# SISTEMI OPERATIVI "OPEN SOURCE": IL CASO LINUX

**Giampaolo Amadori**
**Gianfranco Bazzigaluppi**
**Walter Bernocchi**
**Mauro Gatti**

In questo articolo vengono presentate le caratteristiche di *LINUX* dalla sua storia, all'evoluzione fino all'affermarsi all'interno del più generale fenomeno del software "*Open Source*". Si dimostra come LINUX sia già pronto per il mercato e come permetta la realizzazione di soluzioni di elevata affidabilità, prestazioni, scalabilità, flessibiltà , sicurezza ed economicità. A tale riguardo si analizzano ed approfondiscono le soluzioni basate su *CLUSTER LINUX* che si stanno affermando sul mercato.

## 1. BREVE STORIA DEI SISTEMI OPERATIVI CON PARTICOLARE RIFERIMENTO ALLO UNIX

In origine il termine "software" (SW) veniva usato per identificare quelle parti di un sistema di calcolo che fossero "modificabili" liberamente tramite differenti configurazioni di cavi e spinotti, diversamente dall'"hardware" (HW) con cui si identificava la componentistica elettronica. Successivamente all'introduzione di mezzi linguistici atti ad alterare il comportamento dei computer, SW prese il significato di "espressione in linguaggio convenzionale che descrive e controlla il comportamento della macchina". In particolare occorre notare che ogni macchina aveva un suo proprio linguaggio convenzionale, comunemente detto Assembler.
Nella relativamente breve storia della "Information Technology", tutte le società produttrici di HW hanno cominciato a fornire, assieme all' HW della macchina un certo numero di programmi in grado di svolgere le funzioni di base. Inizialmente si parlava di Monitor, poi di Supervisori, infine è stato adottato il nome di "Sistema Operativo". Tali programmi erano scritti sempre in Assembler, cioè nel linguaggio specifico della singola macchina.
Una eccezione a tale regola è stata il sistema operativo UNIX, realizzato da una organizzazione che non produceva HW, e precisamente i Laboratori Bell, e che pertanto fu progettato fin dall'inizio per risultare indipendente dalla specifica piattaforma HW su cui era stato inizialmente scritto. Non solo, i progettisti iniziali dello UNIX, proprio per renderlo indipendente dalla piattaforma HW, svilupparono un linguaggio, conosciuto come "Linguaggio C", che contiene costrutti della programmazione strutturata (come i cicli FOR e DO...WHILE, le clausole IF...THEN...ELSE, ecc.) oltre a permettere la gestione precisa delle posizioni di memoria, possibile con l'Assembler.
Un'altra peculiarità, è che, avendo allora (anni '70) la Bell una causa in corso per violazione della legge statunitense sui monopoli, lo UNIX veniva inizialmente distribuito corredato dei sorgenti. Questo permise ad altri di portare il sistema UNIX stesso su piattaforme diverse da quelle usate dai laboratori

Bell e ne causò una rapidissima diffusione fra le comunità dei ricercatori e degli sviluppatori. Non solo: la disponibilità dei sorgenti permetteva a chiunque di contribuire al miglioramento ed all'espansione delle funzioni del sistema operativo. In pratica si formò una comunità di sviluppatori volontari, in genere dell'ambiente universitario, ma non solo, che contribuì, in un modo di operare che potremmo definire "collaborativo", alla crescita dello UNIX.
Nel 1984 la Bell perse la causa e la proprietà dei laboratori UNIX passò all'AT&T, che iniziò a rivendicarne i diritti, cioè cominciò a chiedere delle royalty, ma soprattutto mise un freno alla distribuzione dei sorgenti. Ne seguì una battaglia per i diritti di proprietà dello UNIX che durò una decina d'anni. Un folto gruppo di costruttori di sistemi informatici costituirono la Open Software Foundation, con l'obiettivo di farla diventare la proprietaria del software di base comune a tutti e sul quale tutti avrebbero costruito il loro sistema UNIX. Il risultato fu che quell'unico UNIX divenne una molteplicità di Sistemi Operativi proprietari: AIX, Digital UNIX, HP-UX, Sinix, Irix, UNIX386, ecc.

## 1.1. Il progetto GNU e la Free Software Foundation

La principale conseguenza della scomparsa della Bell e della trasformazione di UNIX in tanti sistemi proprietari fu che la comunità di programmatori, formatasi spontaneamente in quegli anni si trovò impossibilitata a continuare a lavorare. Uno degli "hacker" più lungimiranti, Richard M. Stallman, ricercatore presso il Laboratorio di Intelligenza Artificale dell' M.I.T, nel 1985 diede origine al progetto GNU ed alla Free Software Foundation.
Vorrei sottolinerare come l'uso recente della parola "hacker" con il significato di "pirata informatico" è una deformazione dovuta ai mass media. I veri "hacker" rifiutano questo significato e continuano ad usare la parola intendendo "Qualcuno a cui piace programmare e gode nell'essere bravo a farlo" [9].
L'acronimo GNU significa "GNU is Not Unix", ad indicare che l'evoluzione che aveva seguito il sistema operativo UNIX era sbagliata.
La Free Software Foundation (FSF) si occupava e si occupa tuttora di eliminare le restrizioni sulla copia, sulla ridistribuzione, sulla comprensione e sulla modifica dei programmi per computer, concentrandosi in particolare sullo sviluppo di nuovo software libero, inserendolo in un sistema coerente che possa eliminare il bisogno di utilizzare software proprietario.
Stallman sì premurò anche di concepire un'infrastruttura legale entro cui il sistema GNU potesse prosperare. Introdusse la **licenza GPL**, GNU Public License o General Public License, che garantisce le seguenti libertà:
- libertà di eseguire il programma per qualunque scopo, senza limitazioni;
- libertà di adattare il programma alle proprie necessità (l'accesso ai sorgenti è condizione essenziale);
- libertà di copiare il programma senza limitazioni;
- libertà di distribuire le copie modificate (l'accesso ai sorgenti è condizione essenziale).

Oltre a questi diritti, stabilisce un dovere: tutte le modifiche a software licenziato con la GPL devono essere rilasciate con la stessa licenza. La GPL è quindi una licenza persistente.
Il concetto di **Copyleft** - all rights reversed (gioco di parole con Copyright - all rights reserved) è stabilito nelle quattro libertà e nel dovere sanciti nella GPL e che vengono concessi (*left*, participio passato di leave).

## 1.2. Il kernel LINUX

La Free Software Foundation si occupò di riscrivere tutto il sistema UNIX per non dovere royalty a nessuno. Tutti i componenti venivano però innestati su un kernel UNIX, essendo il progetto del Kernel GNU (di nome Hurd) non ancora completato. A tutt'oggi il Kernel Hurd è ancora in fase di sviluppo.
Il 25 agosto 1991 uno studente finlandese, Linus Torvald, annunciò che stava lavorando ad un kernel UNIX-like ("non un progetto professionale come GNU" diceva il suo messaggio) e chiedeva suggerimenti su quali funzioni erano ritenute interessanti; e fornì quindi questo kernel, in seguito ribattezzato LINUX, alla Free Software Foundation.
Il **sistema GNU/Linux** (l'insieme del kernel Linux e delle componenti del progetto GNU) oggi è un sistema operativo completo, funziona su un numero notevolissimo di piat-

taforme hardware. Senza campagne pubblicitarie sponsorizzate da organizzazioni commerciali, ma solo per forza sua, è arrivato in pochi anni ad oltre dieci milioni di installazioni nel mondo. In particolare si stima che Linux sia il motore di oltre sei milioni di server Web in Internet.

## 1.3. Software Libero (Free software) ovvero "Open Source"

Un altro nome per il software libero è "Open Source". Questo termine è nato il 3 febbraio 1998 in una riunione a Palo Alto, California, a cui partecipavano Todd Anderson, Chris Peterson (del Foresight Institute), John "maddog" Hall and Larry Augustin (ambedue di Linux International), Sam Ockman (del Silicon Valley Linux User's Group), and Eric Raymond. Il concetto venne sviluppato durante una discussione sul doppio significato, in inglese, della parola "free": se usata nella frase "free beer tonight", significa "gratis", ma se usato come nel primo emendamento della Costituzione Americana con il concetto di "free speech right" allora significa "diritto alla libertà di parola".
Per evitare questa possibile confusione fu coniato il termine "Open Source". (La storia completa della nascita e della diffusione del termine si può trovare in [2] che, oltre a contenere la storia di OSI, intesa come Open Source Initiative, fornisce anche molti puntatori a documenti che riguardano la storia degli hacker e di Linux.)

### 1.3.1. Il Software Libero è più affidabile e più efficiente

Il vantaggio principale del software libero non consiste, come invece molti vogliono credere, nel fatto che è a basso costo perché si può copiare senza limiti e senza royalty. Il vantaggio principale sta nella sua robustezza. Questa robustezza è dovuta al modo come viene sviluppato il software, che è diverso dal modo di sviluppo del software proprietario. I moduli del software libero vengono sviluppati da poche persone, spesso da una sola persona, ma poi, invece di essere protetti e tenuti nascosti, come accade per il software proprietario, vengono pubblicati in Internet, dove sono a disposizione di tutti i programmatori che desiderano rendersi utili verificando software altrui. In pratica il software viene sottoposto a decine, centinaia di revisioni da parte di professionisti. Eventuali buchi o debolezze vengono evidenziati molto prima che possano emergere durante il funzionamento. Volendo banalizzare si tratta dell'applicazione del vecchio proverbio: "quattro occhi vedono meglio di due", che in questo caso diventa, in inglese, "Many eyes make all bugs shallow".
Per lo stesso motivo il software libero può essere anche più efficiente. Se un hacker, inteso come professionista del software, esaminando una porzione di codice si accorge che esiste un modo più efficiente di risolvere un problema, molto probabilmente lo comunicherà alla comunità degli sviluppatori, magari inviando la parte di codice già scritta. Alla lunga questo meccanismo spontaneo porta a software robusto ed efficiente.
Nel mondo del software libero inoltre il software viene rilasciato senza l'ansia di rispettare le scadenze di annuncio, tipiche del mondo del software proprietario. Non esiste un team di marketing che, avendo già annunciato delle nuove funzionalità, fa pressione sugli sviluppatori perché le date preannunciate siano rispettate. L'obiettivo è mettere a disposizione del software che funzioni bene, non del software che sia disponibile ad una data prefissata. Anzi, lo sviluppatore ha interesse a rilasciare del software che funzioni bene perché non c'è nessun guadagno a distribuire delle correzioni, a differenza di quanto invece avviene con gli "aggiornamenti" di software chiuso e proprietario.

### 1.3.2. Il Software Libero è più sicuro

Nel software libero non possono esistere backdoor, cioè punti di ingresso conosciuti solo da chi ha sviluppato il codice. Essendo i sorgenti a disposizione di tutti, la presenza di una backdoor verrebbe segnalata in tempi brevi. Ovviamente, tutte le volte che si installa del software libero già compilato non ci può essere alcuna garanzia che uno sviluppatore malintenzionato non abbia inserito del codice non conforme ai sorgenti pubblici. Alla lunga però difetti di questo genere vengono scoperti e possono essere corretti, dato che il codice sorgente originale è sem-

pre a disposizione per essere ricompilato.
È accaduto, per esempio, che le installazioni di un noto Data Base Relazionale proprietario contenessero un problema riguardante la sicurezza degli accessi al Data Base, perché banalmente gli sviluppatori avevano lasciato nel codice una backdoor di cui poi si erano dimenticati. Quando il Data Base in questione è diventato libero ed è stato messo a disposizione di tutti in formato sorgente, l'errore è stato scoperto e corretto nel giro di pochi mesi.

#### 1.3.3. Il Software Libero garantisce la scelta del fornitore

Al di là delle questioni tecniche finora elencate, un altro vantaggio tipico del software libero sta nella libertà di scelta: il cliente può davvero liberamente scegliere il fornitore di software o di servizi migliore disponibile sul mercato.
Forse paradossalmente, la massima libertà di scelta favorisce anche i fornitori stessi. Di primo acchito sembra che con il software libero per un cliente sia più semplice abbandonare un fornitore per un altro, ma questo significa anche che il cliente non è spaventato a dover lavorare con una piccola ditta che egli teme possa sparire in cinque o dieci anni. Molti piccoli sviluppatori software temono di perdere i loro già piccoli guadagni se sviluppassero software libero. In realtà molti programmatori e case di sviluppo vivono dignitosamente scrivendo e vendendo software libero, in un mercato in cui le loro scelte sono realmente libere.
Lo stesso non si può dire di chi basa i suoi programmi su piattaforme software proprietarie: devono costantemente subire "scelte di mercato" prese da altri, spendendo moltissima parte di quanto guadagnano solo per aggiornamenti di licenze per software che non fa quanto promesso dalle brochure pubblicitarie.

#### 1.3.4. Il Software Libero abbassa la barriera all'ingresso

Il software libero per sua natura abbassa le barriere di ingresso al mercato e questo può certamente essere salutato come un lieto evento da parte degli sviluppatori e degli utenti. Gli utenti avranno in tale modo la possibilità di ridurre i costi fissi legati alla loro soluzione salvaguardando quindi una maggiore liquidità all'acquisto di servizi per lo sviluppo di nuove soluzioni. Il fiorire di società che stanno migrando o sviluppando applicazioni su Linux così come l'affermarsi di palmari basati sul kernel Linux lo dimostra. Le imprese libere sono compensate con vendite e profitti per aver legalmente e correttamente soddisfatto gli acquirenti. Il software libero in quanto tale in realtà viene venduto a prezzi talmente bassi che i guadagni più significativi vengono fatti vendendo servizi e assistenza o soluzioni. RedHat, Suse, Caldera, TurboLinux e Mandrake sono alcuni esempi di aziende che si guadagnano da vivere con il software libero.
L'alternativa al software libero è la sorveglianza, ovvero assicurarsi che il software non sia usato o copiato illegalmente. Parlando in generale, la parola "sorveglianza" non è usata - si sente invece parlare di "controllo licenze" o frasi simili.

### 1.4. L'affermarsi di LINUX

Oggigiorno tutte le maggiori società di analisi e consulenza del mercato della Information Technology (IT) riconoscono il ruolo di primaria importanza già acquisito da Linux e pure ne prevedono una sempre maggiore affermazione nei prossimi anni. Al riguardo, una delle maggiori società di analisi del mercato IT, nel luglio 2001, sottolineava come LINUX già rappresentasse con ben il 26,9% di diffusione quale OS utilizzato sui nuovi ambienti server installati nel corso dell'anno 2000 il secondo OS più diffuso (Tabella 1).
Le cifre relative alle spedizioni sono, ovviamente, da intendersi in migliaia di unità.
È importante poi notare come il tasso di crescita annuale previsto per Linux sia notevolmente superiore a quello di qualsiasi altro OS e pertanto una sua sempre maggiore diffusione e rilevanza nei prossimi anni risulta cosa certa.
Per quanto riguarda poi le aree applicative per le quali Linux risulta essere maggiormente utilizzato queste sono (in ordine decrescente per quanto concerne la odierna diffusione): l'area dei Web Server, gli e-mail server, i Network server, i Firewall, i Web Application Server, l'area dello sviluppo delle

| Piattaforma SW | 1999 Spedizioni | 1999 Share (%) | 2000 Spedizioni | 2000 Share (%) | 1999-2000 Crescita (%) |
|---|---|---|---|---|---|
| Windows NT/2000 | 2.086 | 38,4 | 2.508 | 409 | 20,2 |
| **Linux** | **1.322** | **24,3** | **1.645** | **26,9** | **24,4** |
| Netware 3.x,4.x,5.x | 1.064 | 16,9 | 1.030 | 16,8 | –3,1 |
| Combined Unix | 826 | 15,2 | 826 | 13,5 | 0 |
| Other NOS | 140 | 2,6 | 116 | 1,9 | –17,4 |
| Total | 5.437 | 100 | 6.125 | 100 | 12,6 |

**TABELLA 1**
*Confronto tra sistemi operativi in funzione del numero di Server installati (Fonte: IDC)*

applicazioni, l'area dei DataBase Server, del Workgroup ed infine quella dei Transaction Server.
Riassumendo, LINUX si è da sempre più diffuso in nuove aree applicative in linea con la sua crescente evoluzione e crescita tecnologica; ovvero, non appena Linux è risultato sufficientemente maturo per essere utilizzato per le applicazioni legate al mondo di internet/intranet è stato da subito usato per esse ed ora che si è ancora più consolidato tecnicamente è pronto per svolgere anche funzioni di DB Server e Transaction Server e pertanto inizia ad essere ora molto usato anche per tali aree applicative.

## 1.5. Fattori che favoriscono e fattori che contrastano il rapido affermarsi di LINUX

Analisi di mercato hanno mostrato come i fattori che maggiormente stanno contribuendo all'affermarsi di Linux siano: il basso costo iniziale, la sua notevole affidabilità e robustezza, il basso costo a regime, il fatto che rappresenti una alternativa a mondi proprietari, la sua sempre crescente scalabilità , la sua portabilità fra piattaforme HW diverse, la crescente disponibilità di applicazioni ed il fatto che sia open source.
Le stesse analisi di mercato hanno invece evidenziato come i fattori che contrastano l'affermarsi di Linux siano: la mancanza di società che ne assicurino servizio e supporto, la mancanza di skill presso i clienti, la mancanza di applicazioni, il fatto che alcune applicazioni ritenute importanti da parte dei clienti debbano essere ancora migrate a tale piattaforma e la mancanza di cultura e conoscenza da parte del mercato.

## 1.6. Le iniziative di IBM atte a supportare l'utilizzo di LINUX

L'IBM ha fatto proprie le analisi e soprattutto le debolezze sintetizzate ai punti precedenti ed avendo riconosciuto in Linux un OS già adatto oggi per un utilizzo sempre più significativo da parte delle aziende ha deciso, ad inizio 2001, di investire un miliardo di dollari in tale area.
Con tale importante investimento l'IBM ha focalizzato le seguenti aree:
- tutte le piattaforme Server sono state rese in grado di supportare Linux offrendo così ai clienti delle "value proposition" ed opportunità rivoluzionarie assolutamente inimmaginabili fino a solo un anno fa;
- è stata intrapresa la migrazione di moltissimi dei prodotti SW IBM (quale DB2, Websphere, Domino, Tivoli, ecc.) su Linux per le diverse piattaforme HW;
- è stato costituito un cosiddetto Linux Technology Center per il quale lavorano a tempo pieno oltre 250 programmatori IBM e che ha la missione di lavorare assieme alla "Linux Community" per accelerare la maturazione di Linux attraverso lo sviluppo di utilities, tools, codice, ecc.;
- il supporto degli "Open Source Development Laboratories" assieme ad altri leader del mercato IT;
- la costituzione di 11 Centri a livello mondiale che sono in grado di aiutare sia Independent Software Vendor (ISV) sia clienti che volessero fare il porting su Linux di applicazioni attualmente utilizzate su altri OS;
- la costituzione ed il mantenimento di diversi siti web ai quali gli utenti possono collegarsi per ricevere una informazione puntualmente aggiornata relativamente al mondo Linux, al-

le applicazioni disponibili, ai tool a disposizione per gli sviluppatori, ecc.;
- ed infine, l'IBM ha iniziato a fornire con proprio personale una notevolissima quantità di Servizi per i clienti quali servizi di supporto di base ed avanzato, servizi di formazione, di consulenza, di implementazione e di sviluppo di applicazioni.

Per maggiori informazioni ci si può riferire al sito web http://www.ibm.com/linux [5].

### 1.7. Il ruolo delle Distribuzioni

Sebbene Linux sia liberamente disponibile e scaricabile attraverso Internet, lo scaricare il codice sorgente ed il ricompilare lo stesso al fine di farlo funzionare correttamente sulla piattaforma HW posseduta non è una attività così semplice da essere alla portata di tutti. Per tale motivo un discreto numero di società ha iniziato a distribuire Linux. Tali società non cambiano il sistema operativo, che è protetto dalla GPL ma si fanno bensì pagare perché portano Linux sul mercato sotto forma di un set consistente di CD, Manuali e pure assicurano il supporto all'installazione ed all'utilizzo di Linux. IBM ha sottoscritto partnership mondiali con le maggiori società che distribuiscono Linux (quali Red Hat, Suse, Caldera e TurboLinux) e pure alcune partnership locali (con Mandrake, RedFlag, ecc.).

### 1.8. LINUX: aree di utilizzo

Al paragrafo 1.4 abbiamo gia velocemente analizzato le aree applicative per le quali Linux risulta essere oggi maggiormente utilizzato. Tali aree applicative possono essere sintetizzate in **"quattro Modelli Infrastrutturali ed Applicativi"** per i quali Linux può costituire una soluzione veramente innovativa.
Questi quattro modelli sono:
- **Appliances**: ovvero i sistemi dedicati allo svolgimento ottimale di una specifica funzione quali i sistemi che fungono da firewall, web-server, web application server, ecc.;
- **Distributed Enterprise**: Linux è un sistema robusto, affidabile, liberamente distribuibile ed economico ed è quindi perfetto per le aziende che hanno decine, centinaia o migliaia di sedi dove debbono utilizzare le stesse applicazioni con sicurezza, controllo ed in maniera economicamente efficace;
- **Linux Cluster**: ovvero insiemi di server che vengono collegati fra loro al fine di realizzare delle soluzioni infrastrutturali in grado di offrire alte prestazioni e/o superiori livelli di sicurezza a condizioni economicamente interessantissime e *dei quali parleremo più in dettaglio anche nei paragrafi successivi*;
- **Workload Consolidation**: ovvero il processo di razionalizzazione e riduzione del fenomeno di selvaggia ed incontrollata proliferazione di decine, centinaia e addirittura migliaia di server che sta oggi colpendo e mettendo in crisi la infrastruttura IT di tante aziende e Service Provider.

Grazie al fatto che oggi Linux può girare eccezionalmente bene ed assai efficacemente anche sui più grossi server IBM delle famiglie zSeries, iSeries e pSeries, e grazie all'estrema possibilità di portare e migrare applicazioni dai mondi proprietari verso il mondo Linux, i clienti possono pensare di "consolidare" centinaia di server su di un solo sistema con enormi vantaggi dal punto di vista del "Total Cost of Ownership" della soluzione, maggiore semplicità gestionale e riduzione delle risorse richieste.

## 2. CLUSTER LINUX: DESCRIZIONE GENERALE

Kai Hwang e Zhiwei Xu nel loro libro *Scalable Parallel Computing* [4], definiscono i *cluster* come insiemi interconnessi di interi computer (*nodi*) che lavorano congiuntamente come un singolo sistema (*single system image*) per fornire un servizio continuativo (*availability*) e servizi efficienti (performance). I cluster il cui scopo prioritario è quello di garantire la disponibilità del servizio sono noti come *HA cluster*, mentre quelli il cui scopo prioritario è di garantire alte prestazioni sono noti come *HPC cluster*. Nel seguito dedicheremo un paragrafo agli HA cluster e uno agli HPC cluster. In entrambi i casi ci limiteremo a considerare i sistemi basati sul sistema operativo Linux.

### 2.1. CLUSTER Linux per la realizzazione di sistemi e soluzioni Altamente Affidabili

Nel disegno di una soluzione ciò che l'architetto deve considerare prioritario, al fine di soddisfare le esigenze dell'azienda che

adotta quella soluzione, è la *realizzazione di un sistema che eroghi il servizio con prestazioni adeguate, nel momento in cui questo servizio effettivamente è richiesto, ad un costo il più piccolo possibile*. *Prestazioni adeguate* significa che, salvo alcune limitate eccezioni, le aziende non sono interessate ad avere il massimo delle prestazioni possibili, ma solo ad avere prestazioni adeguate alle loro esigenze. Per esempio, se il sistema ERP aziendale ha tempi di risposta per l'attività di inserimento ordini dell'ordine di 1 o 2 secondi, ben difficilmente i sistemi informativi saranno disposti ad investire ingenti somme di denaro per dimezzare il tempo di risposta. *Servizio effettivamente richiesto* significa che, anche se i sistemi informativi gradirebbero avere un sistema che non si guasti mai e non richieda alcuna interruzione di servizio per operazioni di manutenzione, questo non è certo l'obiettivo primario.

In effetti gli obiettivi sono, in ordine decrescente di importanza:

1. *minimizzare il costo (danno) dovuto a interruzioni dell'erogazione del servizio*: ciò significa che le interruzioni pianificate di servizio sono considerate accettabili, benché ovviamente non siano auspicabili;
2. *minimizzare il numero e la lunghezza delle interruzioni del servizio*: anche se le interruzioni di servizio pianificate non hanno un impatto grande come quelle non pianificate, nondimeno creano problemi di gestione, hanno comunque un impatto negativo sull'utenza se il sistema lavora 24 ore al giorno, possono ridurre le performance (buffering subottimale) e richiedono il riavvio dei processi di elaborazione batch;
3. *minimizzare il numero dei guasti*: anche se con tecniche di mascheramento è possibile far si che il guasto non produca un'interruzione di servizio, le componenti guasti devono però essere sostituite, con un conseguente incremento di lavoro per i sistemi informativi.

Notiamo che mentre l'obiettivo 1 è di importanza cruciale sia per gli utenti che per i sistemi informativi, l'obiettivo 2 è maggiormente importante per i sistemi informativi e l'obiettivo 3 è esclusivamente rilevante per i sistemi informativi.

Nel seguito del paragrafo faremo vedere come, facendo leva sulle caratteristiche avanzate della piattaforma IBM eServer xSeries e sulla robustezza del sistema operativo Linux, con un accurato disegno della soluzione sia possibile soddisfare la grande maggioranza delle esigenze aziendali per mezzo del sistema operativo Linux installato su server eServer xSeries. Vedremo in particolare come l'uso di tecnologie di clustering permetta di raggiungere livelli di disponibilità del servizio potenzialmente superiori al 99.9%.

### 2.1.1. Evitare i guasti

Per evitare o minimizzare i guasti è necessario che il disegno di tutte le componenti della soluzione, separatamente prese, sia finalizzato all'obiettivo di costruire componenti affidabili, ma anche e soprattutto che nella fase di disegno si tenga esplicitamente conto delle complesse interazioni fra componenti durante l'esecuzione dei processi elaborativi. Quanto detto concerne ovviamente sia le componenti HW che le componenti SW. L'affidabilità del kernel Linux è stata soggetta a diversi studi che ne hanno dimostrato una stabilità paragonabile a quella dei più blasonati sistemi Unix proprietari. L'affidabilità complessiva del sistema operativo Linux su piattaforma Intel, ossia della combinazione del kernel Linux con il SW fornito a corredo dalle distribuzioni e della piattaforma Intel sottostante, è stato oggetto di un recente studio di D.H. Brown. La classificazione di D. H. Brown delle distribuzioni Linux in comparazione ai principali sistemi Unix relativamente al criterio RAS (acronimo che sta per le tre parole Reliability, Availability e Serviceability), ha visto prevalere SuSE Linux 7.2 sulle altre distribuzioni e ha evidenziato alcuni punti deboli di Linux, o più precisamente di Linux su piattaforma Intel. Va detto però che alcuni di questi limiti sono in fase avanzata di soluzione, quale per esempio il multi-path I/O, mentre altri sono in via di soluzione con l'arrivo dei primi server basati sull'IBM Enterprise X-Architecture (EXA).

Il problema estremamente complesso della qualità del SW e dei processi di debugging, verifica e test attraverso i quali si riduce la percentuale di difetti presenti nel SW non sa-

ranno esaminati in questo articolo. Un'analisi approfondita e aggiornata di questo argomento è contenuta nel numero monografico dell'IBM System Journal su *Software Testing and Verification* [3].

### 2.1.2. Evitare le interruzioni di servizio

Evitare o minimizzare le interruzioni di servizio in presenza di guasti è possibile tramite tecniche di *mascheramento*. Un tipico esempio è l'uso di un *fault-tolerant team* di schede di rete. Se una scheda si rompe, o se si rompe lo switch al quale la scheda è connessa, viene attivata, in modo trasparente per gli applicativi, la scheda di backup. Le tecniche di mascheramento sono ampiamente utilizzare nel mondo Intel per dischi (RAID), schede di rete (team fault tolerant), ventole di raffreddamento, alimentatori. Fino a qualche mese fa rimaneva problematico la protezione attraverso ridondanza di CPU, memoria e sottosistema di I/O. La ridondanza nel sottosistema di I/O (multipath I/O) è ora possibile con il nuovo driver delle schede Qlogic. La protezione della memoria sarà possibile a breve con l'arrivo dei primi eServer xSeries che supportano il memory mirroring con sostituibilità a caldo dei banchi di memoria, una delle pietre miliari dell'Enterprise X-Architecture. Rimane come ultimo e più difficile ostacolo la ridondanza con sostituibilità a caldo delle CPU. Attualmente tutto ciò che si riesce a fare in caso di blocco di una CPU è il riavvio del server con successivo isolamento della CPU guasta e ripresa dell'attività sfruttando le altre CPU (in un sistema multiprocessore).

### 2.1.3. Evitare le interruzioni di servizio quando il sistema è utilizzato

La trasformazione del tempo di arresto (downtime) da non pianificato in pianificato è di grande utilità per la grande maggioranza dei sistemi; fanno eccezione ovviamente i sistemi che devono lavorare 24 ore al giorno per 365 giorni all'anno. Fra le tecniche più promettenti per raggiungere questo scopo vale la pena citare la *Software Rejuvenation*. Numerosi studi sviluppati a partire dalle metà degli anni '90 hanno evidenziato che molti blocchi di sistemi sono dovuti a errori di programmazione che provocano durante l'esecuzione dei programmi fenomeni quali *memory leaks, unterminated threads, memory bloating*, ecc. In linea di principio è ovvio aspettarsi che il miglior modo per affrontare questi problemi sia definire migliori tecniche di controllo del SW e, in caso di problema, di identificare la componente malfunzionante per poterla riparare. Vi sono varie ragioni per le quali questa non è, quanto meno, l'unica via per affrontare il problema. Da un lato se il baco risiede in un SW proprietario per ottenere una *patch* che corregga l'errore può essere necessario aspettare mesi, e questa è in effetti una delle ragioni del successo del modello di sviluppo Open Source. D'altro canto alcuni degli errori possono essere così complessi da rendere l'individuazione della causa estremamente difficile (i cosiddetti Heisenbugs). In tutti questi casi è di grande beneficio per il sistema adottare tecniche di Software Rejuvenation, ossia di riavvio controllato di singoli gruppi di processi o di tutto il sistema operativo. Anche se il riavvio controllato è una prassi che molti amministratori di sistema applicano da molto tempo, la determinazione del momento ottimale per effettuare il riavvio e l'individuazione di metodiche che permettano di prevedere il presentarsi di guasti, sono problemi tutt'altro che banali che sono stati oggetto di numerosi studi sviluppati congiuntamente da IBM e dalla Duke University. Sulla base di questi studi, basati su sofisticate tecniche matematiche (*Stochastic Reward Nets*), è stato sviluppato il modulo del SW IBM Director, offerto gratuitamente su tutti gli eServer xSeries, che implementa il processo di SW Rejuvenation.

### 2.1.4. Come gestire i guasti

Se non si riesce ad evitare il guasto e soprattutto se non si riesce ad evitare che il guasto si presenti nel momento in cui sistema è utilizzato, non resta che disporre di metodiche che rendano minimo il tempo che intercorre fra il guasto e la riparazione del sistema. Vari accorgimenti aiutano a ridurre il tempo necessario per la diagnosi e quello necessario per la riparazione. Ma sia la criticità sempre crescente dei sistemi informativi che la complessità degli stessi, stanno spingendo il mercato all'adozione sempre più ampia di meccanismi che automatizzino il processo di

ripristino del servizio. Rientrano in questa categoria in particolare le metodologie di replicazione e quelle di clustering. Le tecniche di replicazione, utilizzate in file system distribuiti come il GPFS di IBM e in tutti i maggiori DBMS ,giocano un ruolo molto importante, ma per contenere l'esposizione non saranno analizzate in questa sede. Affronteremo invece i cluster per alta affidabilità.
I cluster per alta affidabilità possono essere classificati secondo molti punti di vista. Per presentarne le caratteristiche elencheremo nel seguito alcuni di questi punti di vista e faremo vedere come alcuni di questi cluster si comportano relativamente a questi. La trattazione ovviamente non ha alcuna pretesa di esaustività. Il lettore potrà trovare informazioni più complete nel sito http://linux-ha.org [1].

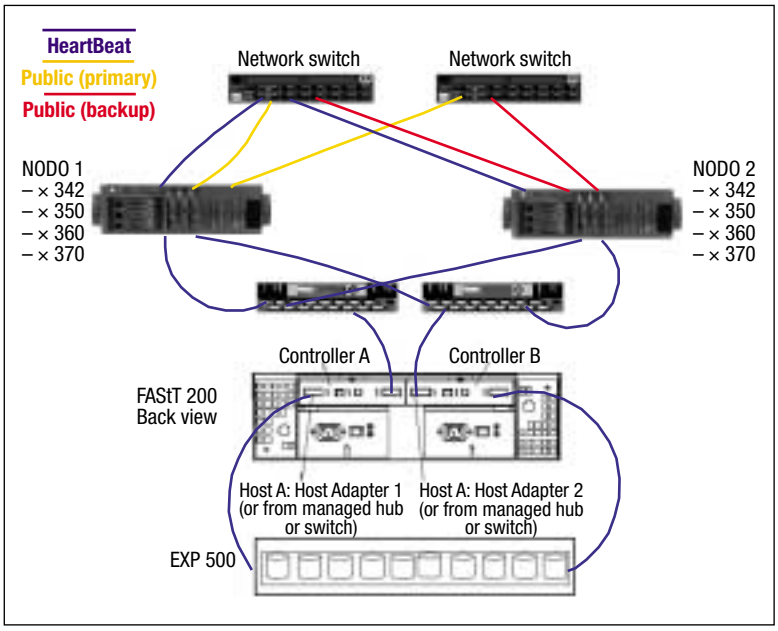



**FIGURA 1**
*Cluster HA linux con IBM eServer xSeries e storage eServer FAStT200*

### L'accesso ai dati

Vi sono sostanzialmente tre categorie di cluster:

- cluster basati sulla totale replicazione dei dati (o clonazione dei server) quali per esempio TurboCluster della TurboLinux o LifeKeeper Data Replication della SteelEye Technologies;
- cluster basati sulla condivisione dei dischi a livello HW ma non a livello SW (*shared nothing*) quali per esempio SteelEye LifeKeeper e Mission Critical DataGuard;
- cluster basati sulla condivisione dei dischi sia a livello HW che ha livello SW (*shared everything*) come per esempio Oracle 9i Real Application Cluster.

La prima categoria di cluster è prevalentemente utilizzata per le web farm o per altre attività a basso carico transazionale. La seconda categoria di cluster è quella con il più ampio spettro di applicazioni. Infatti, può essere utilizzata per proteggere virtualmente ogni applicativo: web server, database server, application server, ecc. La terza categoria di cluster richiede lo sviluppo di un complesso meccanismo per evitare danni all'integrità dei dati dovuti all'accesso simultaneo di più nodi, ossia di un distributed lock manager; per questa ragione non ha mai raggiunto un elevato livello di diffusione. Nel seguito ci concentreremo prevalentemente sulla seconda categoria.

### Il sottosistema dischi condiviso

I nodi possono accedere al sottosistema dischi condiviso per mezzo del protocollo SCSI, del protocollo FC (SAN) o attraverso il protocollo IP (NAS). Di questi il protocollo FC è quello che garantisce la migliore combinazione di affidabilità e performance. Per questa ragione la grande maggioranza dei cluster per alta affidabilità, quantomeno per i server basati su piattaforma Intel, utilizzano la tecnologia FC. I NAS sono un'alternativa interessante per sistemi che non abbiano le esigenze di performance e integrità dei dati che sono tipiche dei sistemi transazionali. La figura 1 rappresenta un cluster Linux per alta affidabilità basato su IBM eServer xSeries e storage IBM FAStT200.

### Le distribuzioni

Un aspetto fondamentale è l'integrazione con le distribuzioni. La tabella 2 fornisce un elenco sintetico delle distribuzioni supportate.

### Gli applicativi

In linea di principio un cluster per alta affidabilità shared nothing può proteggere ogni applicativo "che si comporti bene" ossia che sia in grado di ripristinare la propria attività in seguito ad un'interruzione irregolare. Per esempio un database transazionale può essere protetto con questa tecnica perché il DBMS riporta il sistema in condizioni di inte-

**TABELLA 2**
*Distribuzione supportata dai Software*

| SW | Distribuzioni supportate |
|---|---|
| SteelEye LifeKeeper v. 4.0 | Red Hat 6.x and 7.x, SuSE 7.x, Caldera eServer 2.3, Miracle Linux 1.1 and 2.x e TurboLinux (vedi http://www.steeleye.com/products/linux/system_conf.html per un elenco aggiornato) |
| Mission Critical Convolo Data Guard | Red Hat Linux 6.2, 7.1, SuSE Linux 7.1 e Debian GNU/Linux 2.2r3 (vedi http://www.mclinux.com/products/convolo-requirements.php per un elenco aggiornato) |

grità sfruttando i log. Accanto al problema di quali applicativi possono essere protetti vi è il problema di come proteggere gli applicativi. In generale i SW di clustering forniscono meccanismi generici per controllare lo stato dei processi, spegnerli in caso di malfunzionamento e riavviarli su un altro nodo. L'implementazione specifica dei meccanismi di controllo, avvio e spegnimento dei processi è però demandata ad opportuni script. SteelEye Technologies vende degli Application Recovery Kit che possono essere utilizzati per agevolare il processo di configurazione di applicativi in cluster e un Software Development Kit per consentire lo sviluppo di opportuni script per applicativi per i quali non è disponibile un Application Recovery Kit. L'elenco degli ARK disponibili presso SteelEye Technologies è contenuto nella seguente pagina web http://www.steeleye.com/products/supported_apps.html [7]. Mission Critical fornisce invece supporto per un limitato numero di applicazioni sotto forma di servizio. L'elenco degli applicativi supportati è disponibile all'URL http://www.missioncriticallinux.com/support/supported-applications.php [8].

### 2.1.5. Un Esempio di Linux Cluster per la High Availability (HA)

IBM e Open4.it hanno realizzato un cluster HA e Load Balanced, basato su architettura Pentium, fiber channel e software OpenSource per Linux con le seguenti caratteristiche:

- 40 cpu pentium 4;
- 35 GB ram;
- 1 Tera di storage.

Il cluster è articolato in: una coppia di bastioni in HA, una coppia di balancer in HA, venti server in LB come farm, una coppia di database server in HA, una coppia di publishing server in HA, una coppia di file server in HA, uno storage in fiber channel con controller ridondato, due switch fiber channel e due switch Catalist (Cisco) di rete.
Tale cluster è attuamente in funzione presso Aonet in Milano, dove assicura la continuità dei servizi di web publishing, content management, ecommerce, messaging, groupware web e wap che vengono offerti in modalità ASP.
In particolare le applicazioni che vengono fatte girare sono: Apache, Interchange, Qmail, Courier, Twig, Zope, NFS e vario software aggiuntivo per sicurezza e intrusion detection
Secondo Maruzzelli (Resp. Architettura per Open4.it) "Le ragioni che hanno portato all'acquisto di tale cluster sono il vantaggio competitivo che si realizza da una installazione con cui riescono ad assicurare performance, sicurezza e continuità di servizio ad un gran numero di clienti, con costi decrescenti e con amministrazione unificata. Il vantaggio specifico dell'OpenSource e di Linux risiede invece nell'ampia comunità di sviluppatori software e amministratori che garantiscono un costante aggiornamento delle applicazioni sia in termini di funzionalità che, aspetto molto importante, di sicurezza".

## 2.2. CLUSTER LINUX per la realizzazione di sistemi di High Performance Computing(HPC)

I termini High Performance Computing e High Throughput Computing denotano un ampio gruppo di sistemi caratterizzati da grandi esigenze elaborative. Un elenco anche solo sommario dei settori che fanno uso di cluster per HPC richiederebbe molte pagine. Citiamo a puro titolo di esempio:

- fisica delle alte energie - per esempio il pro-

getto LHC del CERN per la ricerca dell'evanescente bosone di Higgs e la verifica delle teorie supersimmetriche;
- prospezioni geologiche - particolarmente rilevante per le aziende del settore petrolifero;
- studi geologici, in particolare per il problema della previsione dei terremoti;
- Life Sciences - sia la genomica che la proteomica richiedono enormi potenze elaborative per analizzare Petabytes di dati;
- previsioni meteorologiche - per analizzare modelli sempre più dettagliati dell'atmosfera;
- Data mining - per analizzare gli enormi database generati dai sistemi di CRM al fine di evitare di perdere clienti e garantire un livello di servizio più elevato;
- risk management - ad esempio per minimizzare i rischi operativi nei settori finanziari e lungo la supply chain;
- analisi di portafoglio - particolarmente interessante per le società di intermediazione mobiliare che facciano uso di algoritmi statistici o di intelligenza artificiale al fine di prevedere l'andamento dei titoli e/o contenere il rischio di investimenti.

In moltissimi settori nei quali si fa uso di algoritmi per i quali la latenza della comunicazione fra i processi non è critica (*coarse grain parallelism*), i cluster sono il metodo preferenziale per la gestione di elaborazioni complesse. In particolare i cluster basati su Linux si stanno affermando come piattaforma preferenziale per l'High Throughput Computing. Le principali ragioni alla base di questa evoluzione sono:
- il basso costo dei server basati su piattaforma Intel (per esempio gli IBM eServer xSeries);
- l'accresciuta performance erogata dai server Intel, in particolare per i calcoli su interi e più recentemente con il rilascio dei Pentium IV anche per i calcoli in virgola mobile;
- il basso costo del sistema operativo Linux;
- la disponibilità del codice sorgente del sistema operativo Linux che permette di ottimizzarlo in funzione dell'esigenza; citiamo a titolo d'esempio la creazione di sistemi operativi con processi mobili (MOSIX, Qlusters OS);
- la facilità con la quale gli esperti gestori dei sistemi HPC, che hanno prevalentemente una cultura UNIX, possono apprendere l'uso di un sistema Linux;
- le ottime caratteristiche del kernel Linux, particolarmente a partire dal rilascio dei kernel della famiglia 2.4.x.

Notiamo altresì che i tentativi fatti di utilizzare Windows NT/2000 in luogo di Linux non hanno avuto molto successo. Le principali ragioni sono:
- costo del sistema operativo;
- difficoltà di ottimizzazione data l'indisponibilità del codice sorgente;
- sistema di gestione meno flessibile.

### 2.2.1. Beowulf: una delle architetture dei Linux cluster

Introduciamo ora una delle architetture maggiormente utilizzate per la realizzazione di Linux cluster per HPC, dalla definizione di T. Sterling e Don Becker in "How to Build a Beowulf": "*un Beowulf è un gruppo di personal computers (PCs), interconnessi da una qualsiasi tecnologia di rete, su cui girano uno dei sistemi operativi open source con caratteristiche proprie dello Unix*". A titolo esemplificativo, un cluster Beowulf potrebbe essere composto da semplicissimi PC, uno switch di rete che riconosca il protocollo TCP/IP e Linux come sistema operativo. Chiunque abbia una discreta conoscenza della tecnologia Intel, del sistema operativo Linux, e del protocollo TCP/IP può realizzare un Linux cluster.

I server e l'infrastruttura di rete sono però, per quanto importanti, solo due delle componenti di un cluster Linux per HPC. Occorre infatti in primo lugo un applicativo "parallelo", ossia in grado di sfruttare il parallelismo del sistema HW. Inoltre occorre un sistema di gestione efficiente, poiché fino a quando il numero di nodi è piccolo, la manutenzione cluster non è molto difficoltosa, ma quando il numero dei nodi diviene elevato il sistema diventa rapidamente ingestibile in mancanza di un sistema gestione efficiente. In particolare in virtù del progressivo aumento del numero di guasti al crescere del numero di nodi. A tal proposito esistono vari tool sia commerciali che non commerciali che aiutano i system administrator del cluster, per una descrizione di questi tools e delle componenti che fanno parte di un Linux cluster per HPC rimandiamo alla lettura dell'IBM redbook SG24-6041-00 scaricabile all'indirizzo http://www.redbooks.ibm.com [6].

### 2.2.2. Le componenti di Linux cluster per HPC

In termini generali un cluster Linux per High Performance Computing è costituito da

- nodi di elaborazione (per esempio eServer xSeries 330);
- un architettura di rete ad alte prestazioni per la comunicazione fra i nodi (per esempio basata su schede di rete e switch Myrinet) indicata spesso con l'espressione *system area network*;
- un sistema per l'immagazzinamento dei dati;
- sistema operativo Linux (spesso basato sulla distribuzione RedHat);
- compilatori paralleli (Portland Group Fortran F90, C, C++, GNU compilers, Intel F90 compiler);
- librerie parallele (per esempio MPI o PVM);
- Job Management System;
- un sistema di gestione integrato del cluster.

Il tipici nodi utilizzati in un cluster Linux per HPC sono

- server monoprocessori basati sull'architettura IA32 (per esempio eServer xSeries 330);
- server biprocessori basati sull'architettura IA32 (per esempio eServer xSeries 330);
- server quadriprocessori basati sull'architettura IA64 (per esempio eServer xSeries 380).

La scelta del tipo di server dipende essenzialmente dal tipo di applicativo, da considerazioni di costo e da valutazioni sulla complessità di gestione. Notiamo altresì che mentre Linux supporta teoricamente fino a un massimo di 16 processori, il grado di scalabilità al di sopra di 8 processori è tutt'altro soddisfacente con gli attuali kernel. Viceversa, i sistemi basati sull'architettura IA64 rendono possibile l'indirizzamento di grandi quantità di RAM, ma sono attualmente limitati ad un massimo di 4 processori. Un altro punto che occorre tenere presente è che, in un cluster costituito da molti nodi, una percentuale significativa del costo complessivo è rappresentata dai dispositivi a corredo (rack, cablaggio, monitor, ecc.). Al fine di minimizzare questa componente sono utili tecnologie quali l'IBM C2T technology che può consentire una riduzione dei costi dell'ordine di decine di migliaia di dollari.

La *system area network* può essere basata su schede di rete e switch Fast Ethernet, se gli applicativi possono tollerare un'alta latenza nella comunicazione fra processi, da schede e switch Gigabit Ethernet, se è necessaria una più grande larghezza di banda o infine da dispositivi Myrinet, se è fondamentale minimizzare la latenza nella comunicazione fra processi. È importante sottolineare in proposito che in un cluster costituito da decine di nodi l'adozione di un'infrastruttura di rete ad alte prestazioni può far si che il costo della rete eguagli o addirittura superi quello dei server. Una chiara comprensione delle effettive esigenze elaborative è quindi cruciale per contenere i costi.

Il sistema per l'immagazzinamento dei dati può consistere semplicemente dei dischi contenuti nei server o alternativamente di uno o più file server. In questo secondo caso un ruolo particolarmente cruciale è svolto non solo dal tipo di file server, per esempio un IBM Enterprise Storage Server o piuttosto un IBM Network Attached Storage, ma anche dal tipo di cluster file system con il quale si accede a questi dati (per esempio l'IBM General Parallel File System).

Un Job Management System è costituito da tre parti principali:

- uno *user server* che consente agli utenti di sottomettere i job a uno o più code di elaborazione;
- un *job scheduler* che decide come e quando assegnare le risorse ai processi di elaborazione;
- un *resource manager* che effettua l'allocazione delle risorse e controlla l'esecuzione dei job.

Il sistema per la gestione integrata del cluster deve garantire per quanto possibile

- una gestione integrata del cluster che consenta di soddisfare il requisito della *Single System Image*;
- l'installazione automatizzata, controllata e a blocchi della server farm;
- garantire un'elevata fault tolerance.

Per quanto concerne le problematiche di gestione ci limitiamo a citare il prodotto open source xCAT sviluppato da IBM, nonché il blasonato IBM Cluster Management System, recentemente portato su Linux. Relativamente alla fault tolerance è evidente che mentre è cruciale per un cluster costituito da centinaia di nodi minimizzare il costo per nodo ed è spesso accettabile in ca-

so di rottura di un nodo il dover riavviare i job in esecuzione su quel nodo, non è però al tempo stesso accettabile avere un'alta difettosità. Se per esempio un job richiede settimane per essere eseguito, il doverlo rilanciare comporta un ritardo di settimane. Tecniche di Predictive Failure Analysis quali quelli disponibili sugli eServer xSeries possono essere d'aiuto per raggiungere questo obiettivo.

### 2.2.3. Un'alternativa avanzata

Accanto alle metodologie tradizionali che prevedono lo sviluppo di applicativi paralleli per mezzo di librerie quali PVM e MPI si sta diffondendo in vari settori l'alternativa più avanzata di utilizzare la migrazione di processi per ottimizzare la distribuzione del carico fra i nodi di elaborazione. Particolarmente interessante nel mondo Linux è il MOSIX sviluppato presso l'Università di Gerusalemme dal Prof. Amnon Barak del quale è stata recentemente presentata una versione commerciale (Qlusters OS) dall'azienda israeliana Qlusters Inc. Quest'ultimo prodotto contiene due importanti componenti che mancano nella versione open source (MOSIX): uno scheduler e un queue manager.

### 2.2.4. Un Esempio di Linux Cluster per HPC

Un esempio di Linux cluster per HPC in Italia è quello che IBM ha realizzato per il Cineca di Bologna e che ha le caratteristiche riportate in tabella 3.

Tale cluster è oggi in funzione presso il Cineca di Bologna dove è utilizzato per fare girare codici di calcolo relativi alle seguenti aree : astrofisica, chimica computazionale e fisica dello stato solido.

Le ragioni per le quali il Cineca ha deciso di installare un tale HPC Cluster sono dovute alla convinzione che sia gia oggi utile sperimentare le tecnologie di clustering Linux e Beowulf per il supercalcolo a conferma della crescente maturità raggiunta da Linux.

| | |
|---|---|
| Compute node Model: | IBM x330 servers |
| Processor (PE): | Intel Pentium III 1.133 GHz |
| Number of PEs: | 128 |
| Processor per node: | 64 nodes with 2 proc. |
| L2 Cache: | 512 kbit per processor full speed |
| DRAM: | 64 Gbytes (1GB/node) |
| Disk space: | 2.7 TB |
| Peak performance: | 145 Gflop/s |
| OS: | Linux |
| Internal Network: | Miricom LAN Card "C" Version |
| Available compilers: | Portland Group Fortran F90, C, C++<br>GNU compilers<br>Intel F90 compiler |
| Parallel libraries: | MPI |

**TABELLA 3**
*Caratteristiche di un cluster HPC*

## Ringraziamenti

Giunti al termine di tale articolo ci fa piacere il ricordare e ringraziare il Dr. Stefano Maffulli (Associazione Culturale MILUG di Milano) per il prezioso supporto fornitoci e per la passione con la quale promuove il software Open Source, il Dr. Erbacci (CINECA di Bologna) per il supporto offerto per quanto riguarda le informazioni relative alla installazione CINECA ed il Dr. Maruzzelli per le informazioni e considerazioni forniteci relativamente alla installazione Open4.it.

Vogliamo poi concludere con una esortazione ed un incitamento al guardare già oggi a LINUX come una possibilità reale che può offrire un vantaggio competitivo all'interno della vostra azienda, del vostro istituto, della vostra scuola perchè sicuramente vi sono gia oggi moltissime aree della vostra infrastruttura informatica che potrebbero trarre dei vantaggi, a volte impensabili, dall'utilizzo di Linux e di altro codice Open Source.

## Bibliografia

[1] High-Availability Linux Project: HTTP://WWW.LINUX-HA.ORG

[2] History of the OSI: http://www.opensource.org/docs/history.html

[3] IBM: Software Testing and Verification. *IBM System Journal*, Vol. 41, n. 1, 2002. http://www.research.ibm.com/journal/sj41-1.html

[4] Kai Hwang, Zhiwei Xu: *Scalable Parallel Computing: Technology,Architecture,Programming*. McGraw-Hill Companies, 1998. http://shop.barnesandnoble.com/booksearch/results.asp?WRD=hwang+xu&userid=66GToH2D8T

[5] Linux at IBM: http://www-1.ibm.com/linux

[6] RedBooks: http://www.redbooks.ibm.com

[7] Supported Apps: http://www.steeleye.com/products/supported_apps.html

[8] Supported Layered Products: http://www.missioncriticallinux.com/support/supported-applications.php

[9] The GNU Project: http://www.gnu.org/gnu/the-gnu-project.html

GIAMPAOLO AMADORI ingegnere nucleare con specializzazione successiva in marketing, è responsabile IBM delle attività di Vendita e Marketing Linux per la Regione Sud Europea. In IBM, dal 1989, ha ricoperto diversi ruoli di sempre maggiore responsabilità quali: Responsabile attività di Marketing per il Mercato Italiano delle Small and Medium Businesses, Direttore Vendite Italia per i Midrange Server.
E-mail: giampaolo_amadori@it.ibm.com

GIANFRANCO BAZZIGALUPPI è Manager of University Relations di IBM Italia.
Esperto di economia e organizzazione, attivo nella ricerca (IreR; Centro Studi di IBM) e nella docenza universitaria, dal 1994 al febbraio 2001 è stato responsabile della Fondazione IBM Italia.
Giornalista pubblicista, ha diretto dal 1993 al 1999 la rivista If, edita da G. Mondadori.
E-mail: bazzi@it.ibm.com

WALTER BERNOCCHI è un Advisor IT Specialist e lavora in IBM Italia per il gruppo xSeries Pre-Sale Technical Support. Ha iniziato la sua collaborazione con IBM nel 1989. Da marzo 2000 lavora a tempo pieno per sviluppare e supportare soluzioni in ambiente Linux. Ha esperienza di C/C++, Java, SunOS, Solaris e AIX. E' Red Hat Certified Engineer (RHCE).
E-mail: walter_bernocchi@it.ibm.com

MAURO GATTI dottore in Fisica e in Ricerca in Fisica, è team leader dell'IBM EMEA eServer xSeries Solution Architects team. La sua area di specializzazione è il disegno infrastrutturale di e-business server farms con eServer xSeries. In particolare si occupa di progetti di sistemi altamente affidabili, Linux, Grid Computing e ERP. Ha esperienze come free lance per alcune grandi aziende dell'IT (Microsoft, HP, IBM).
E-mail: mauro_gatti@it.ibm.com

# COSCIENZA ARTIFICIALE: MISSIONE IMPOSSIBILE?

**Giorgio Buttazzo**

Potrà mai un computer diventare autocosciente? La coscienza è una prerogativa degli esseri umani o potrebbe svilupparsi in un sistema artificiale altamente complesso? Dipende dal materiale di cui sono fatti i neuroni oppure può essere replicata su un hardware differente? Più che fornire delle risposte a tali interrogativi, questo articolo cerca di porre nuove domande, identificare i problemi e discutere le possibili implicazioni legate allo sviluppo di un sistema artificiale autocosciente.

## 1. INTRODUZIONE

Fin dai primi sviluppi dell'informatica, i ricercatori hanno speculato sulla possibilità di costruire macchine intelligenti capaci di competere con l'uomo. Oggi, considerati i successi dell'intelligenza artificiale e la velocità di sviluppo dei computer, tale possibilità sembra essere più concreta che mai, tanto che molti cominciano a credere che in un futuro non tanto lontano le macchine raggiungeranno e supereranno l'intelligenza umana, fino a sviluppare una mente cosciente. Quando si parla di coscienza artificiale, tuttavia, occorre anche considerare gli aspetti filosofici della questione. Un processo di calcolo che evolve in un computer può essere considerato simile al pensiero? Viceversa, il pensiero di una mente umana può essere considerato un processo di calcolo? La coscienza è una prerogativa dei soli esseri umani? Dipende dal materiale di cui sono fatti i neuroni o può essere replicata su un hardware differente?
Oggi nessuno è in grado di fornire una risposta esauriente a questi interrogativi, e allo stato attuale delle conoscenze possiamo solo imbastire una discussione di carattere interdisciplinare a cavallo di settori quali l'informatica, la neurofisiologia, la filosofia e la religione. Anche la fantascienza, essendo un prodotto dell'immaginazione umana, ha un ruolo importante in questa discussione, in quanto, esprimendo i desideri e le paure umane sulla tecnologia, può influenzare il corso del progresso. In effetti, a livello sociale, la fantascienza agisce fornendo una sorta di simulazione di scenari futuri che contribuiscono a preparare la gente ad affrontare transizioni cruciali o a prevedere le conseguenze dell'uso di nuove tecnologie.

## 2. LA VISIONE FANTASCIENTIFICA

Nella letteratura e nella cinematografia fantascientifica, le macchine pensanti costituiscono una presenza costante. Sin dai primi anni cinquanta, i film di fantascienza hanno ritratto i robot come macchine sofisticate costruite per svolgere operazioni complesse al servizio dell'uomo, per lavorare in ambienti

ostili, oppure pilotare astronavi in viaggi galattici [3]. Nello stesso tempo, però, i robot intelligenti sono stati spesso ritratti anche come macchine pericolose, capaci di cospirare contro l'uomo attraverso piani subdoli. L'esempio più significativo di robot con queste caratteristiche è HAL 9000, protagonista principale del film *2001 Odissea nello Spazio* di Stanley Kubrick (1968). Nel film, HAL controlla l'intera nave spaziale, parla dolcemente con gli astronauti, gioca a scacchi, fornisce giudizi estetici su quadri, riconosce le emozioni dei membri dell'equipaggio, ma finisce per uccidere quattro dei cinque astronauti per perseguire un piano elaborato al di fuori degli schemi programmati [14].
In altri film di fantascienza, come T*erminator e Matrix*, la visione del futuro è ancora più catastrofica: i robot diventano autocoscienti e prendono il sopravvento sulla razza umana. Ad esempio, *Terminator 2 – Il giorno del Giudizio*, di James Cameron (1991), comincia col mostrare una guerra terribile tra robot e umani, mentre sullo schermo compare la scritta:

LOS ANGELES, 2029 A.D.

Secondo la trama del film, dopo che la Cyberdyne era diventata la maggiore ditta fornitrice di sistemi militari intelligenti computerizzati, tutti i sistemi di difesa erano stati modificati per diventare completamente autonomi e supervisionati da Skynet, un potente processore neurale costruito per prendere decisioni strategiche di difesa. Tuttavia, Skynet comincia ad apprendere con una velocità esponenziale e diventa autocosciente 25 giorni dopo la sua attivazione. Presi dal panico, gli ingegneri provano a disattivarlo, ma Skynet reagisce violentemente scatenando un attacco militare contro gli umani [4].
Solo in pochissimi film, i robot sono descritti come degli assistenti affidabili, che cooperano con l'uomo piuttosto che cospirare contro. Nel film *Ultimatum alla Terra*, di Robert Wise (1951), Gort è forse il primo robot (extraterrestre in questo caso) che affianca il capitano Klaatu nella sua missione di portare il messaggio di pace agli umani al fine di evitare la loro autodistruzione. Anche in *Aliens - Scontro Finale* (secondo episodio della fortunata serie, diretto da James Cameron nel 1986), Bishop è un androide sintetico il cui scopo è di pilotare l'astronave durante la missione e proteggere l'equipaggio. Rispetto al suo predecessore (presente nel primo episodio), Bishop non è affetto da malfunzionamenti e rimane fedele alla sua missione fino alla fine del film.
Tra gli altri robot buoni ricordiamo Andrew, il robot maggiordomo NDR-114 nel film *L'uomo bicentenario* (Chris Columbus, 1999) e David, il robot bambino nel film *A.I. Intelligenza Artificiale* (Steven Spielberg, 2001). Entrambi cominciano ad operare come macchine preprogrammate, ma col tempo finiscono per acquisire una coscienza di esistere.
La duplice connotazione spesso attribuita ai robot della fantascienza rappresenta il desiderio e la paura che l'uomo ha verso la propria tecnologia. Da una parte, infatti, l'uomo proietta nel robot il suo irrefrenabile desiderio d'immortalità, materializzato in un essere artificiale potente e indistruttibile, le cui capacità intellettive, sensoriali e motorie sono amplificate rispetto a quelle di un uomo normale. D'altro canto, però, esiste la paura che una tecnologia troppo avanzata (quasi misteriosa per la maggior parte di persone) possa sfuggire di mano e agire contro lo stesso uomo (vedi la creatura del Dr. Frankenstein, HAL 9000, Terminator, e i robot di *Matrix*). Il cervello positronico dei robot di Isaac Asimov deriva dalla stessa sensazione di disagio: esso era il risultato di una tecnologia così sofisticata che, sebbene il processo per costruirlo fosse completamente automatizzato, nessuno conosceva più i dettagli del suo funzionamento [2].
I recenti progressi dell'informatica hanno influenzato le caratteristiche dei robot della fantascienza moderna. Ad esempio, le teorie sul connessionismo e le reti neuronali artificiali (mirate a replicare i meccanismi di elaborazione tipici del cervello umano) hanno ispirato il robot Terminator, il quale non è solo intelligente, ma è in grado di apprendere in base alla sua esperienza. Terminator rappresenta il prototipo del robot che abbiamo sempre immaginato, in grado di camminare, parlare, percepire il mondo, tanto da essere indistinguibile da un essere umano. Le sue batterie

possono fornirgli energia per 120 anni, e un circuito di alimentazione alternativo consente di tollerare i danni al circuito primario. Ma, la cosa più importante è che Terminator può apprendere! Egli è controllato da un processore neurale, ossia un computer che può modificare il suo comportamento in base alle esperienze vissute.
Dal punto di vista filosofico, l'aspetto più intrigante sollevato dal film, è che tale processore neurale è così complesso che comincia ad apprendere a velocità esponenziale fino a diventare autocosciente! In tal senso, il film solleva una questione importante sulla coscienza:

"*Può mai una macchina diventare autocosciente?*"

Prima di affrontare questa questione, tuttavia, dovremmo chiederci:

"*Come possiamo verificare che un essere intelligente sia autocosciente?*".

## 3. IL TEST DI TURING

Nel 1950, il pioniere dell'informatica Alan Turing si pose un problema simile, ma riguardo all'intelligenza. Allo scopo di stabilire se una macchina possa o no essere considerata intelligente come un essere umano, egli propose un famoso test, noto come il test di Turing: un computer e una persona interagiscono con un esaminatore esterno attraverso due terminali. L'esaminatore pone delle domande su qualsiasi argomento utilizzando una tastiera. Sia il computer che l'uomo inviano delle risposte che l'esaminatore legge sui monitor corrispondenti. Se l'esaminatore non è in grado di determinare a quale terminale è connessa la persona e a quale il computer, si dice che il computer ha superato il test di Turing e può essere considerato intelligente come l'umano che è dall'altra parte.
Nel 1990, il test di Turing ha ricevuto il suo primo riconoscimento formale da parte di Hugh Loebner e del Cambridge Center for Behavioral Studies del Massachusetts, che hanno instaurato il premio Loebner in Intelligenza Artificiale [11]. Loebner offre un premio di 100.000 dollari per il primo computer in grado di fornire risposte indistinguibili da quelle di un umano. La prima competizione è stata tenuta presso il Computer Museum di Boston nel Novembre del 1991. Per qualche anno, la gara è stata limitata ad un singolo argomento ristretto, ma le competizioni più recenti, a partire dal 1998, hanno eliminato la restrizione sugli argomenti del colloquio. Ciascun giudice, dopo la conversazione, da un punteggio da 1 a 10 per valutare l'interlocutore, dove 1 significa umano e 10 computer [9]. Finora, nessun computer ha fornito risposte totalmente indistinguibili da un umano, ma ogni anno i punteggi medi ottenuti dai computer tendono ad essere sempre più vicini a 5. Oggi, il test di Turing può essere superato da un computer solo se si restringe l'interazione su argomenti moto specifici, come gli scacchi.
L'11 Maggio 1997, per la prima volta nella storia, un computer chiamato Deep Blue ha battuto il campione del mondo di scacchi Garry Kasparov, per 3.5 a 2.5. Come tutti i computer, tuttavia, Deep Blue non comprende il gioco degli scacchi come potrebbe fare un umano, ma semplicemente applica delle regole per trovare la mossa che porta ad una posizione migliore, in base ad un criterio di valutazione programmato da scacchisti esperti.
Claude Shannon ha stimato che nel gioco degli scacchi lo spazio di ricerca della soluzione vincente comprende circa 10120 possibili posizioni. Deep Blue era in grado di analizzare circa 200 milioni ($2 \cdot 10^8$) di posizioni al secondo. Per esplorare l'intero spazio di ricerca Deep Blue impiegherebbe circa $5 \cdot 10^{111}$ secondi, corrispondenti a $10^{95}$ miliardi di anni, un tempo di gran lunga superiore all'età del nostro universo (stimata sui 15 miliardi di anni). Ciononostante, la vittoria di Deep Blue è attribuibile alla sua velocità di calcolo combinata con gli algoritmi di ricerca utilizzati, in grado di considerare vantaggi posizionali e materiali (è stato stimato che Kasparov elabori le mosse ad una velocità di tre posizioni al secondo). In altre parole, in questo caso la superiorità del computer è dovuta all'applicazione della forza bruta, piuttosto che ad un'intelligenza artificiale sofisticata.

Nonostante ciò, in molte interviste rilasciate alla stampa, Garry Kasparov ha rivelato che in certe situazioni aveva la sensazione di giocare non contro una macchina, ma contro un umano. Spesso ha apprezzato la bellezza di certe soluzioni ideate dalla macchina, come se essa fosse guidata da un piano intenzionale, piuttosto che da una forza bruta. Dunque, se accettiamo la visione di Turing, dobbiamo ammettere che Deep Blue ha superato il test e gioca a scacchi in modo intelligente.
Oltre agli scacchi, vi sono altri settori in cui i computer stanno raggiungendo le capacità umane, e il loro numero aumenta ogni anno. Nella musica, per esempio, esistono molti programmi commerciali in grado di generare linee melodiche, o addirittura interi brani, secondo stili desiderati, che vanno da Bach alla musica jazz. Esistono anche programmi in grado di generare improvvisazioni eccellenti su basi armoniche predefinite, secondo stili di grandi jazzisti come Charlie Parker e Miles Davis, molto meglio di quanto possa fare un musicista umano di medio livello. Nel 1997, Steve Larson, un professore di musica dell'università dell'Oregon, propose una variante musicale del Test di Turing, chiedendo ad una commissione di ascoltare un insieme di brani di musica classica e di indicare quali fossero stati scritti da un computer e quali da un compositore autentico. Il risultato fu che molti brani generati dal computer furono classificati come composizioni autentiche e viceversa, indicando che anche in questo campo il computer ha superato il Test di Turing [10].
Altri settori in cui i computer stanno diventando abili come gli umani includono il riconoscimento del parlato, la diagnosi degli elettrocardiogrammi, la dimostrazione di teoremi, ed il pilotaggio di velivoli. Nel prossimo futuro, tali settori si espanderanno velocemente e comprenderanno attività sempre più complesse, come la guida di automobili, la traduzione simultanea di lingue, la chirurgia medica, la tosatura dell'erba nei giardini, la pulizia della casa, la sorveglianza di aree protette, fino ad includere forme artistiche, finora prerogativa dell'uomo, come la pittura, la scultura e la danza.
Tuttavia, anche ammesso che le macchine diventino abili come l'uomo, o più dell'uomo, in molte discipline, tanto da essere indistinguibili da esso nel senso indicato da Turing, potremmo concludere che esse hanno raggiunto l'autocoscienza? Naturalmente no. Qui la questione diventa delicata, in quanto, se l'intelligenza è l'espressione di un comportamento esteriore che può essere osservato e misurato mediante test specifici, la coscienza di un individuo è una proprietà interna del cervello, un'esperienza soggettiva, che non può essere misurata dall'esterno. Al fine di affrontare questo problema è necessario svolgere alcune considerazioni filosofiche.

## 4. LA VISIONE FILOSOFICA

Da un punto di vista puramente filosofico, non è possibile verificare la presenza di una coscienza in un altro cervello (sia esso umano che artificiale), in quanto questa è una proprietà che può essere osservata solo dal suo possessore. Poiché non è possibile entrare nella mente di un altro essere, allora non potremo mai essere sicuri della sua capacità di essere cosciente. Tale problema è affrontato in modo approfondito da Douglas Hofstadter e Daniel Dennett in un libro intitolato *L'Io della Mente* [8].
Da un punto di vista più pragmatico, comunque, potremmo seguire l'approccio di Turing e dire che una creatura può essere considerata autocosciente se è capace di convincerci, superando alcune prove specifiche. Inoltre, tra gli uomini, la credenza che un'altra persona sia autocosciente è anche fondata su considerazioni di similitudine: poiché tutti noi siamo fatti allo stesso modo è ragionevole credere che la persona che ci sta davanti sia anch'essa autocosciente. Chi metterebbe in dubbio la coscienza del proprio migliore amico? Se invece la creatura di fronte a noi, pur avendo sembianze e comportamenti umani, fosse fatta di tessuti sintetici, organi meccatronici e cervello a microprocessore neurale, forse la nostra conclusione sarebbe diversa.
L'obiezione più comune che spesso viene mossa contro la coscienza artificiale è che i computer, poiché sono azionati da circuiti elettronici che funzionano in modo automatico e deterministico, non possono essere

creativi, provare emozioni, amore, o avere libero arbitrio. Un computer è uno schiavo azionato dai suoi componenti, proprio come una lavatrice, anche se più sofisticato. Il punto debole di questo ragionamento, tuttavia, è che esso può essere applicato anche alla controparte biologica. Infatti, al livello neurale, il cervello umano è anch'esso attivato da reazioni elettrochimiche e ogni neurone risponde automaticamente ai suoi segnali di ingresso in base a precise leggi fisiche. Nonostante ciò, questo modo di operare del cervello non ci preclude la possibilità di provare felicità, amore, ed avere comportamenti irrazionali.

Con lo sviluppo delle reti neurali artificiali, il problema della coscienza artificiale è diventato ancora più intrigante, in quanto le reti neurali tendono a replicare il funzionamento di base del cervello, fornendo un supporto appropriato per realizzare dei meccanismi di elaborazione simili a quelli che operano nel cervello. Nel libro *Impossible Minds*, Igor Aleksander [1] affronta questo argomento con profondità e rigore scientifico.

Se si rimuove la diversità strutturale tra cervello biologico e artificiale, la questione sulla coscienza artificiale può solo diventare di tipo religioso. In altre parole, se si crede che la coscienza umana sia determinata da un intervento divino (diretto o indiretto), allora chiaramente nessuna macchina potrà mai diventare autocosciente. Se invece si crede che la coscienza umana sia una naturale proprietà elettrica sviluppata dai cervelli complessi, allora la possibilità di realizzare un essere artificiale cosciente rimane aperta.

## 5. LA VISIONE RELIGIOSA

Da un punto di vista religioso, l'argomento principale che viene portato contro la possibilità di replicare l'autocoscienza in un essere artificiale è che la coscienza non è una conseguenza dell'attività elettrochimica del cervello, ma un'entità immateriale separata, spesso identificata con l'anima. Questa visione dualistica del problema mente-corpo fu sviluppata principalmente da Cartesio (1596-1650) ed è ancora condivisa da molte persone. Tuttavia, essa ha perso di credibilità nella comunità scientifica e filosofica, poiché presenta diversi problemi che non possono essere spiegati con tale teoria.

- Innanzitutto, se una mente è separata dal cervello di cui fa parte, come può fisicamente interagire con il corpo e attivare un circuito neurale? Quando io penso di muovermi, nel mio cervello avvengono specifiche reazioni elettrochimiche che fanno sì che alcuni neuroni si attivino per attuare i muscoli desiderati. Ma se la mente opera al di fuori del cervello, in che modo essa è in grado di muovere gli atomi per creare impulsi elettrici? Esistono delle forze misteriose (non contemplate dalla fisica) che attivano le cellule neurali? Dobbiamo ammettere che la mente interagisce con il cervello violando le leggi fondamentali della fisica?
- Secondo, se una mente cosciente può esistere al di fuori del cervello, perché noi abbiamo un cervello?
- Terzo, se le emozioni e i pensieri vengono da fuori, perché un cervello stimolato per mezzo di elettrodi e droghe reagisce generando pensieri?
- Infine, perché in pazienti affetti da malattie cerebrali il comportamento cosciente viene seriamente sconvolto dalla rimozione chirurgica di alcune aree del cervello?

Questi ed altri argomenti hanno portato il *dualismo* a perdere credibilità nella comunità scientifica e filosofica. Per risolvere queste inconsistenze, sono state sviluppate molte altre formulazioni alternative sulla questione mente/corpo. Da un lato estremo, il *riduzionismo* rifiuta di riconoscere l'esistenza della mente come esperienza soggettiva privata, considerando tutte le attività mentali come stati neurali specifici del cervello. All'altro estremo, l'*idealismo* rifiuta l'esistenza del mondo fisico, considerando tutti gli eventi sensoriali come il prodotto di costruzioni mentali. Purtroppo, in questo articolo non c'è spazio per discutere esaurientemente tutte le varie teorie sulla questione, e comunque ciò esula dagli scopi di questo lavoro. Tuttavia, è importante notare che, con il progresso dell'informatica e dell'intelligenza artificiale, gli scienziati e i filosofi hanno formulato una nuova interpretazione, in base alla quale la mente è considerata come una forma di calcolo che emerge ad un livello di astrazione più alto rispetto a quello dell'attività neuronale.

## 6. LA VISIONE OLISTICA

Il maggior difetto dell'approccio riduzionistico nella comprensione della mente è di decomporre ricorsivamente un sistema complesso in sottosistemi più semplici, fino ad arrivare ad analizzare e descrivere le unità elementari. Questo metodo funziona perfettamente per i sistemi lineari, in cui le uscite possono essere viste come la somma di componenti semplici. Tuttavia, un sistema complesso è spesso non lineare, pertanto l'analisi delle sue componenti elementari non è sufficiente a comprendere il suo funzionamento globale. In tali sistemi, ci sono delle caratteristiche olistiche che non possono apparire ad un livello inferiore di dettaglio, ma che si manifestano solo quando si considera la struttura globale e le interazioni fra le diverse componenti.
Paul Davies, nel suo libro *Dio e la nuova fisica* [5], spiega questo concetto osservando che una fotografia digitale di un volto è composta da un grande numero di punti colorati (pixel), ciascuno dei quali non rappresenta il volto: l'immagine prende forma solo quando si osserva la foto da una certa distanza che ci permette di vedere tutti i pixel. In altre parole, il volto non è una proprietà dei singoli pixel, ma solo dell'insieme dei pixel.
In *Göedel, Escher, Bach* [7], Douglas Hofstadter espone lo stesso concetto descrivendo il comportamento di una colonia di formiche. Come si sa, le formiche mostrano una struttura sociale complessa e altamente organizzata basata sulla distribuzione del lavoro e sulla responsabilità collettiva. Sebbene ciascuna formica abbia intelligenza e capacità limitate, l'intera colonia mostra un comportamento estremamente complesso. Infatti, la costruzione di un formicaio richiede un progetto sofisticato, ma chiaramente, nessuna formica ha in mente il quadro completo dell'intero progetto. Ciononostante, al livello di colonia emerge un comportamento complesso e finalizzato. In un certo senso, il formicaio può essere considerato un essere vivente.
Sotto diversi aspetti, un cervello può essere paragonato ad un formicaio, poiché composto da miliardi di neuroni che cooperano per raggiungere un obiettivo comune. L'interazione tra neuroni è molto più stretta che tra le formiche, ma i principi di base sono simili: suddivisione del lavoro e responsabilità collettiva. La coscienza non è una proprietà dei neuroni individuali, i quali operano automaticamente come degli interruttori, rispondendo ai segnali di ingresso in base a precise leggi fisiche; ma è piuttosto una proprietà olistica che emerge dalla cooperazione fra neuroni quando il sistema raggiunge un livello di complessità sufficientemente organizzato.
Sebbene la maggior parte di persone non ha problemi ad accettare l'esistenza di caratteristiche olistiche, qualcuno rifiuta di credere che una coscienza possa emergere da un substrato di silicio, ritenendo (non si sa su quale base) che essa sia una proprietà intrinseca dei materiali biologici, come le cellule neurali. Dunque è ragionevole porci la seguente domanda:

"*Può la coscienza dipendere dal materiale di cui sono fatti i neuroni*?"

Paul and Cox, in *Beyond Humanity* [13] dicono: "Sarebbe sbalorditivo se il più potente strumento di elaborazione delle informazioni potesse essere costruito solo a partire da cellule organiche e chimiche. Gli aerei sono fatti con materiali diversi da quelli di cui sono composti gli uccelli, i pipistrelli, e gli insetti; i pannelli solari sono costruiti con materiali diversi da quelli che costituiscono le foglie. Di solito esiste più di un modo per costruire un dato tipo di macchina. [...] Il materiale organico è solo quello con cui la genetica è stata in grado di lavorare. [...] Altri elementi, o combinazioni di elementi, possono rivelarsi migliori allo scopo di elaborare informazione in modo autocosciente."
Se, dunque, sosteniamo l'ipotesi che la coscienza sia una proprietà olistica del cervello, allora la successiva domanda da porci è:

"*Quando un computer diventerà autocosciente?*"

## 7. UNA PREVISIONE AZZARDATA

Tentare di fornire una risposta, se pur approssimativa, alla precedente domanda è senza dubbio azzardato. Ciononostante è possibile determinare almeno una condizione necessaria, senza la quale una macchina non può sviluppare autocoscienza. L'idea si basa sulla semplice considerazione che, per sviluppare

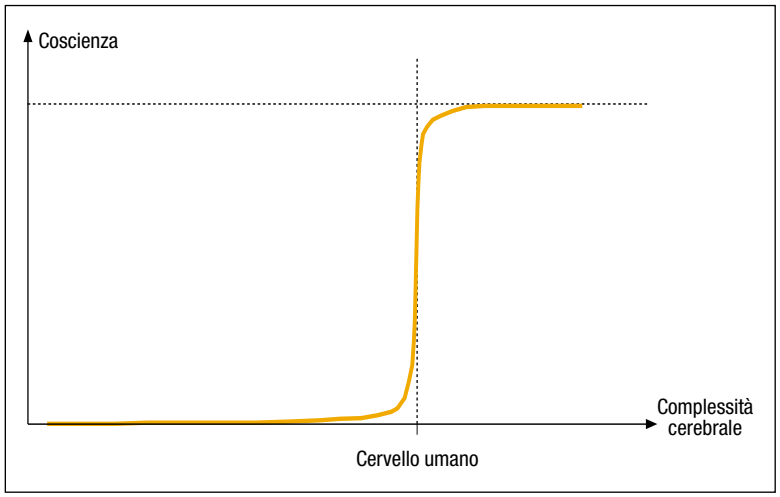



**FIGURA 1**
*Coscienza come funzione della complessità cerebrale*

una forma di autocoscienza, una rete neurale deve essere complessa almeno quanto il cervello umano.
Questa è un'ipotesi ragionevole da cui partire, poiché sembra che i cervelli animali, meno complessi di quello umano, non siano in grado di produrre pensieri consci, nel senso inteso in questo articolo. Dunque, l'autocoscienza sembra essere una funzione della complessità cerebrale a forma di gradino, in cui la soglia è rappresentata dalla complessità del cervello umano (Figura 1).
Ma quanto è complesso il cervello umano? Quanta memoria è richiesta per simulare il suo comportamento con un computer? Il cervello umano è composto da circa mille miliardi ($10^{12}$) di neuroni, e ciascun neurone forma mediamente mille ($10^3$) connessioni (sinapsi) con gli altri neuroni, per un totale di $10^{15}$ sinapsi. In una rete neurale artificiale una sinapsi può essere efficacemente simulata attraverso un numero reale, che richiede 4 byte di memoria per essere rappresentato in un computer. Di conseguenza, per simulare $10^{15}$ sinapsi occorre una memoria di almeno $4\_10^{15}$ byte (4 milioni di Gbyte). Possiamo dunque ritenere che, tenendo conto delle variabili ausiliarie per memorizzare lo stato dei neuroni e altri stati cerebrali, per simulare l'intero cervello umano siano necessari circa 5 milioni di Gbyte. Allora la questione diventa:

"*Quando sarà disponibile tanta memoria in un computer?*"

Durante gli ultimi 20 anni, la capacità della memoria RAM è aumentata in modo esponenziale di un fattore 10 ogni 4 anni. Il grafico di figura 2 illustra ad esempio la tipica configurazione di memoria installata su un personal computer dal 1980 al 2000.
Per interpolazione, si può facilmente ricavare la seguente equazione, che fornisce la dimensione di memoria RAM (in byte) in funzione dell'anno:

$$\text{bytes} = 10^{\left(\frac{\text{year}-1966}{4}\right)}$$

Per esempio, utilizzando l'equazione possiamo dire che nel 1990 un personal computer era tipicamente equipaggiato con 1 Mbyte di RAM. Nel 1998, una tipica configurazione possedeva 100 Mbyte di RAM, e così via. Invertendo la relazione precedente, è possibile predire l'anno in cui un computer sarà dotato di una certa quantità di memoria (assumendo che la RAM continuerà a crescere con la stessa velocità):

$$year = 1966 + 4 \log_{10} (\text{bytes})$$

Duunque, per conoscere l'anno in cui un computer possiederà 5 milioni di Gbyte di RAM, dovremo sostituire tale numero nell'equazione precedente e calcolare il risultato. La risposta è:

**year = 2029**

Un interessante coincidenza con la data predetta nel film Terminator! È interessante anche osservare che una simile previsione è stata derivata indipendentemente da altri studiosi, come Hans Moravec [12], Ray Kurzweil [10], Gregory Paul and Earl Cox [13].
Allo scopo di comprendere a fondo il significato del risultato ottenuto, è importante fare alcune considerazioni. Innanzitutto, è bene ricordare che la data è stata calcolata sulla base di una condizione necessaria, ma non sufficiente, allo sviluppo di una coscienza artificiale. Ciò significa che l'esistenza di un potente computer dotato di milioni di gigabyte di RAM non è sufficiente da solo a garantire che esso diventerà magicamente autocosciente. Ci sono altri fattori importanti che influiscono su questo processo, quali il progresso delle teorie sulle reti neurali artificiali e la compren-

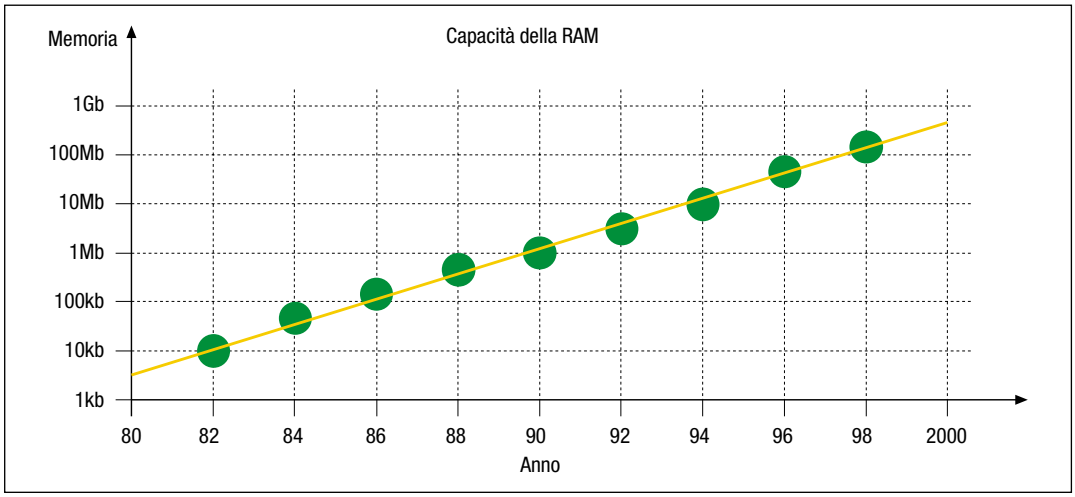



**FIGURA 2**
*Tipica configurazione della RAM (in byte) installata sui personal computer negli ultimi venti anni*

sione dei meccanismi biologici del cervello, sui quali è impossibile tentare una stima precisa. Inoltre, qualcuno potrebbe obiettare che il calcolo effettuato si riferisce alla memoria di un personal computer, che non rappresenta il top della tecnologia nel settore. Qualcun altro potrebbe osservare che la stessa quantità di memoria RAM potrebbe essere disponibile utilizzando una rete di computer oppure dei meccanismi di memoria virtuale che sfruttino lo spazio di uno o più hard disk. In ogni caso, anche utilizzando diversi numeri, il principio di base rimarrebbe lo stesso, e la data verrebbe anticipata solo di qualche anno.

### 7.1. Cosa dire della legge di Moore?

Alcuni obiettano che la previsione del 2029 si basi su una cieca estrapolazione della tendenza attuale di crescita della RAM, senza considerare gli eventi che potrebbero alterare questa tendenza. La tendenza di crescita esponenziale della potenza di calcolo dei computer fu notata nel 1973 da Gordon Moore, uno dei fondatori dell'Intel, il quale predisse che *il numero di transistor nei circuiti integrati sarebbe continuato a raddoppiare ogni 18 mesi fino al raggiungimento dei limiti fisici*. L'accuratezza di tale previsione negli ultimi 25 anni è stata tale da riferire tale osservazione come la "Legge di Moore". Ma fino a quando tale legge continuerà a valere? Le maggiori industrie di circuiti integrati hanno stimato che la Legge di Moore continuerà ad essere valida per altri 15 o 20 anni. Dopo, quando i transistor raggiungeranno la dimensioni di pochi atomi, l'approccio tradizionale non funzionerà più e il paradigma di costruzione dei circuiti sarà destinato a cambiare. Cosa succederà dopo? Forse l'evoluzione dei microprocessori conoscerà la fine intorno all'anno 2020?

Alcuni studiosi, come Ray Kurzweil [10] e Hans Moravec [12], hanno notato che la crescita esponenziale della potenza di calcolo nei computer è cominciata molto prima dell'invenzione dei circuiti integrati, nel 1958, indipendentemente dall'hardware utilizzato. Di conseguenza, la Legge di Moore sui circuiti integrati non descrive il primo paradigma evolutivo, ma il quinto. Ciascun nuovo paradigma ha sostituito il precedente quando necessario. Ciò suggerisce che la crescita esponenziale non si fermerà con la fine della Legge di Moore. L'industria non è a corto di nuove idee per il futuro e si stanno già sperimentando nuove tecnologie, quali il progetto di chip tridimensionali, i computer ottici e i computer quantistici [6]. Dunque, sebbene la Legge di Moore non sarà più valida nel futuro (poiché non si può applicare ai sistemi non basati sul silicio), la crescita esponenziale della potenza di calcolo dei computer probabilmente continuerà per molti anni a venire.

## 8. ALTRI ASPETTI

### 8.1. È preclusa la coscienza alle macchine sequenziali?

Se la coscienza è una proprietà olistica del cervello emergente dal funzionamento collettivo di una struttura neurale altamente

organizzata, e se tale proprietà non dipende dal materiale con cui sono fatti in neuroni, ma dal tipo di elaborazione che essi svolgono, allora è abbastanza ragionevole credere che la coscienza potrebbe anche emergere in una rete neurale artificiale avente una complessità paragonabile o superiore a quella di un cervello umano. Tuttavia, cosa possiamo dire sulla coscienza in computer sequenziali?".

"*Potrebbe mai una macchina sequenziale sviluppare una coscienza?*"

Se la coscienza è il prodotto dell'elaborazione delle informazioni in un sistema altamente organizzato, allora essa non può dipendere dal particolare substrato hardware che realizza il supporto di calcolo. Di fatto, la maggior parte delle reti neurali oggi utilizzate non sono realizzate a livello hardware, ma simulate in un computer sequenziale, che è molto più flessibile (sebbene più lento) di una rete hardware. Qualcuno potrebbe osservare che una simulazione di un processo è diversa dal processo stesso. Chiaramente ciò è vero quando si simula un fenomeno fisico, quale un uragano o un sistema planetario. Tuttavia, per una rete neurale, la simulazione non è diversa dal processo, poiché entrambi sono dei sistemi di elaborazione delle informazioni. Analogamente, la calcolatrice software disponibile sui nostri PC effettua le stesse operazioni della sua controparte elettronica. Dunque, una calcolatrice hardware e la sua simulazione sequenziale sono funzionalmente equivalenti. Pertanto dobbiamo concludere che se una coscienza può emergere da una rete neurale hardware, essa deve necessariamente svilupparsi anche in una sua simulazione software.

## 8.2. La cognizione del tempo

Può la coscienza dipendere dalla velocità degli elementi di calcolo? Non è facile rispondere a questa domanda, ma l'intuizione suggerisce che dovrebbe essere indipendente, in quanto i risultati di un calcolo non dipendono dall'hardware su cui essi sono eseguiti. Ciononostante, la velocità di calcolo è importante per soddisfare i requisiti temporali imposti dal mondo esterno. Se potessimo idealmente rallentare i nostri neuroni uniformemente in tutto il cervello, probabilmente percepiremmo il mondo come in un film accelerato, in cui gli eventi si susseguono ad una velocità maggiore di quella delle nostre capacità reattive. Ciò è ragionevole, in quanto il nostro cervello si è evoluto ed adattato in un mondo in cui gli eventi importanti per la riproduzione e la sopravvivenza sono dell'ordine delle centinaia di millisecondi. Se potessimo accelerare gli eventi dell'ambiente oppure potessimo rallentare i nostri neuroni, non riusciremmo più ad operare efficacemente in tale mondo, e probabilmente non sopravviveremmo a lungo.
Viceversa, cosa proveremmo se avessimo un cervello più veloce? Oggi, una porta logica è un milione di volte più veloce di un neurone: infatti, i neuroni biologici rispondono entro alcuni millisecondi ($10^{-3}$ s), mentre i circuiti elettronici hanno dei tempi di risposta dell'ordine dei nanosecondi ($10^{-9}$ s). Questa osservazione porta ad una domanda interessante: se mai una coscienza emergerà in una macchina artificiale, quale sarà la percezione del tempo in un cervello milioni di volte più veloce di quello umano?
È ragionevole supporre che ad una macchina cosciente il mondo attorno a sé sembri evolversi più lentamente, come un film al rallentatore. Forse la stessa cosa accade agli insetti, che possiedono un cervello più piccolo ma più reattivo e veloce del nostro. Forse, agli occhi di una mosca, una mano umana che tenta di colpirla appare muoversi lentamente, dando ad essa tutto il tempo di volare via comodamente.
Paul and Cox affrontano questa questione nel libro *Beyond Humanity* [13]: " ... un essere cibernetico sarà in grado di apprendere e pensare a velocità elevatissime. Essi osserveranno un proiettile sparato da una distanza moderata, calcoleranno la sua traiettoria e lo eviteranno se necessario. [...] Immaginiamo di essere un robot vicino ad una finestra. Alla nostra mente veloce, un uccello che attraversi il nostro campo visivo sembra impiegare delle ore, e un giorno sembra durare in eterno." È interessante notare che il problema della percezione del tempo nelle menti cibernetiche è stato considerato per la prima volta

nel cinema nel film *Matrix* (Larry & Andy Wachowski, 1999).

## 9. CONCLUSIONI

Dopo tante discussioni sulla coscienza artificiale, viene da porsi un'ultima domanda:

> "*Perché dovremmo costruire una macchina cosciente?*"

A parte problemi etici, che potrebbero influenzare significativamente il progresso in questo campo, la motivazione più forte verrebbe certamente dall'innato desiderio dell'uomo di scoprire nuovi orizzonti ed allargare le frontiere della scienza. Inoltre, lo sviluppo di un cervello artificiale basato sugli stessi principi di funzionamento di quello biologico fornirebbe un modo per trasferire la nostra mente su un supporto più veloce e robusto, aprendo una via verso l'immortalità. Liberati da un corpo fragile e degradabile, e dotati di organi sintetici sostituibili (incluso il cervello), i nuovi esseri rappresenterebbero il successivo passo evolutivo della razza umana. Questa nuova specie, risultato naturale del progresso tecnologico umano, avrebbe la possibilità di esplorare l'universo, sopravvivere alla morte del sistema solare, cercare altre civiltà aliene, controllare l'energia dei buchi neri, e muoversi alla velocità della luce trasmettendo le informazioni necessarie per replicarsi su altri pianeti.
L'esplorazione dello spazio finalizzata alla ricerca di civiltà aliene intelligenti è già cominciata nel 1972, quando la sonda Pioneer 10 fu lanciata per abbandonare il sistema solare con lo scopo preciso di trasmettere nello spazio informazioni sulla razza umana e il pianeta Terra, come un messaggio in una bottiglia nell'oceano.
Tuttavia, come per tutte le più importanti scoperte dell'uomo, dall'energia nucleare alla bomba atomica, dall'ingegneria genetica alla clonazione umana, il problema reale è stato e sarà quello di tenere la tecnologia sotto controllo, assicurando che essa venga usata per il progresso dell'umanità, e non per scopi catastrofici. In questo senso, il messaggio portato dal capitano Klaatu nel film "Ultimatum alla Terra" rimane ancora il più attuale!

## Bibliografia


[1] Igor Aleksander: *Impossible Minds: My Neurons, My Consciousness*. World Scientific Publishers, October 1997.

[2] Isaac Asimov: *I, Robot* (a collection of short stories originally published between 1940 and 1950). Grafton Books, London, 1968.

[3] Giorgio Buttazzo: *Can a Machine Ever Become Self-aware?. In Artificial Humans*, an historical retrospective of the Berlin International Film Festival 2000, Edited by R. Aurich, W. Jacobsen and G. Jatho, Goethe Institute, Los Angeles, May 2000, p. 45-49.

[4] James Cameron, William Wisher: *Terminator 2*: Judgment Day. The movie script, http://www.stud.ifi.uio.no/~haakonhj/Terminator/Scripts/, 1991.

[5] Paul Davis: *God and the New Physics*. Simon and Schuster Trade, 1984.

[6] Linda Geppert: Quantum transistors: toward nanoelectronics. *IEEE Spectrum*, Vol. 37, n. 9, September 2000.

[7] Douglas Hofstadter: *Göedel, Escher, Bach. Basic.* Books, New York, 1979.

[8] Douglas R. Hofstadter, Daniel C. Dennett: *The Mind's I*. Harvester/Basic Books, New York 1981.

[9] Marina Krol: Have We Witnesses a Real-Life Turing Test?. *IEEE Computer*, Vol. 32, n. 3, March 1999, p. 27-30.

[10] Ray Kurzweil: *The Age of Spiritual Machines*. Viking, 1999, p. 160.

[11] Home page of the Loebner Prize, 1999 (Current 28 January 1999): http://www.loebner.net/Prizef/ loebner-prize.html.

[12] Hans Moravec: *Robot: Mere Machine to Transcendent Mind*. Oxford University Press, 1999.

[13] Gregory S. Paul, Earl Cox: *Beyond Humanity: CyberEvolution and Future Minds*. Charles River Media, Inc., 1996.

[14] David G. Stork: *HAL's Legacy: 2001's computer as dream and reality*. Edited by David G. Stork, Foreword by Arthur C. Clarke, MIT Press, 1997.



GIORGIO BUTTAZZO è Professore Associato di Ingegneria Informatica presso l'Università di Pavia, dove svolge attività di ricerca nei settori dei sistemi in tempo reale, della robotica avanzata e delle reti neurali artificiali. Nel 1987, ha conseguito un Master in Computer Science presso l'Università della Pennsylvania e nel 1991 il Dottorato di Ricerca presso la Scuola Superiore S. Anna di Pisa.
E-mail: buttazzo@unipv.it

# WEB LEARNING: ESPERIENZE, MODELLI E TECNOLOGIE

**Alberto Colorni**

Nell'articolo vengono presentate alcune regole per orientarsi nello sviluppo (ma anche nella valutazione) di progetti di e-learning, viene proposto un sistema di classificazione basato su tre elementi, vengono fornite alcune indicazioni sulle tecnologie e sugli strumenti più consueti. Il lavoro nasce dall'esperienza dell'autore nel campo dell'e-learning, sia nella veste di docente al Politecnico di Milano che in quella di direttore del centro per l'innovazione didattica dello stesso Ateneo. L'articolo si conclude con una breve disamina delle principali obiezioni sulla formazione on-line e con la proposta di alcune idee per superarle.

## 1. SCUOLA PER CORRISPONDENZA O ALTRO ?

Il concetto di spartiacque mi ha sempre affascinato: due gocce di pioggia che sono inizialmente distanti meno di un millimetro finiscono l'una nell'Oceano Indiano e l'altra nel Mar Baltico. Può accadere qualcosa del genere anche per l'e-learning: con poche differenze si potrebbe favorire un ulteriore (e definitivo?) degrado dell'istruzione o promuovere un'opportunità nuova e addirittura stimolante.
In questo articolo cercherò di spiegare le ragioni per cui è possibile e sensato fare dell'online learning senza troppe rinunce. Lo farò avendo in mente alcune obiezioni che in questi anni ho sentito spesso fare e che sono esposte molto bene (insieme a molte altre, su cui peraltro concordo) da C. Stoll [18]: obiezioni sulla mancanza di contatto diretto, sui costi sproporzionati, sul pericolo di semplificazione, sull'appiattimento, in una parola su tutto ciò che si perderebbe con l'adozione di questi mezzi.
Vorrei peraltro evitare il trionfalismo di chi pensa a questi strumenti come a una grande rivoluzione che ci renderà tutti pronti per un futuro "digitale" in cui la potenziale libertà di accesso a milioni di dati produrrà istruzione e cultura (e non, come è invece possibile, un faida-te curricolare che ne è l'esatto opposto).
Ciò che voglio (di)mostrare lo riassumo fin d'ora. Si tratta di poche cose, quasi dei consigli per gli acquisti:

- è opportuno farsi guidare dalle esigenze degli utenti e non dalle tecnologie;
- è necessario lavorare (e molto) sull'interattività e sui vari tipi di feedback;
- è fondamentale progettare un percorso, che offra oltre ai materiali anche dei servizi, delle modalità d'interazione, una responsabilità precisa nella conduzione.

Sviluppo le mie considerazioni da un punto di osservazione privilegiato: sono infatti il direttore del Centro METID del Politecnico di Milano [5] che ha prodotto (o partecipato a) molte esperienze nel settore e sono anche il coordinatore del primo progetto di laurea on-line italiano [12].
Per ragioni di spazio userò un approccio abbastanza schematico, rimandando eventuali

ulteriori approfondimenti a un dibattito successivo, magari on-line.
Inizio quindi subito proponendo una distinzione, che è anche indicativa di tre fasi storiche, tra:
**i** distance learning, in cui si è affrontato l'elemento territoriale;
**ii** e-learning, in cui si è inserita la tecnologia informatica;
**iii** on-line learning (o web learning), in cui diventa importante l'interattività.
Naturalmente ogni modalità comprende le precedenti. In questo lavoro mi occupo (quasi esclusivamente) dell'ultima.

## 2. AI VERTICI DI UN CUBO

Normalmente nella Formazione a Distanza (FaD) si segnalano le rotture di continuità rispetto allo spazio e al tempo. Vorrei qui introdurre un terzo elemento su cui ragionare: la modalità di relazione tra i soggetti coinvolti.
Definisco così i tre assi di uno "spazio della formazione", all'interno del quale possono trovar posto alcuni modelli didattici, più o meno noti, di FaDoL (Formazione a Distanza on-line), come mostrato in figura 1. Considerando una coppia di opzioni rispetto a ciascun asse, si ottiene la situazione che segue:

- spazio → presenza (locale) vs. distanza (remoto);
- tempo → reale (sincrono) vs. differito (asincrono);
- relazione → gerarchica (broadcast) vs. collaborativa (interazione, rete).

In ognuno dei tre casi, la seconda opzione introduce una questione rilevante: rispetto allo spazio la questione del canale di trasmissione, cioè delle tecnologie; rispetto al tempo la questione del ritmo di apprendimento e quindi del monitoraggio; rispetto alla relazione la questione dei ruoli ma soprattutto degli attori del processo formativo.
Seguendo la numerazione della figura, illustro sinteticamente le varie situazioni ottenute combinando le opzioni sopra esposte.

Caso 1: è la ***didattica classica***, basata sull'unità di luogo e di tempo e su un solo vero attore (il docente); le versioni più recenti di questo modello offrono supporti come collegamenti Internet e visione di CD-rom ma non ne cambiano la sostanza.

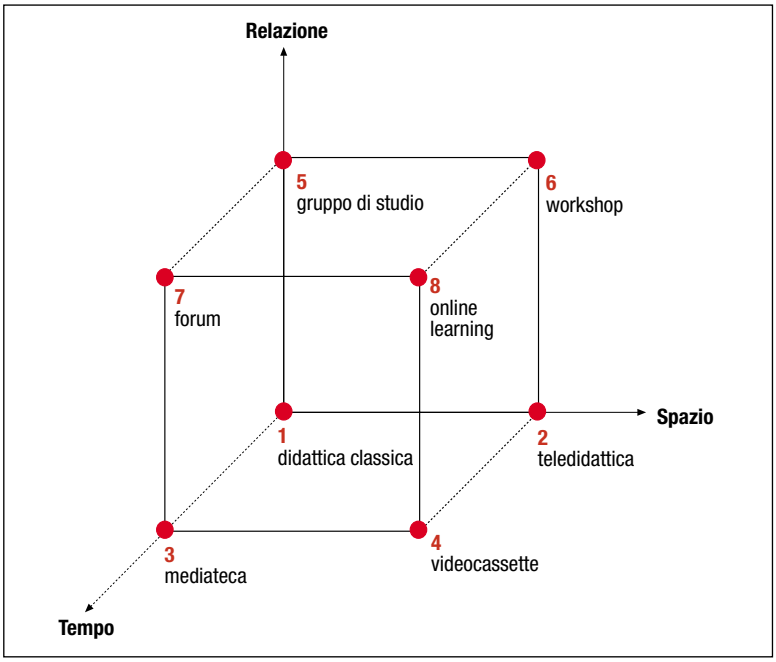



**FIGURA 1**
*Lo "spazio della formazione"*

Caso 2: è la situazione della ***teledidattica*** (e prima ancora della didattica televisiva), con una o più aule remote collegate in tempo reale con l'aula principale dove opera il docente; al Politecnico corsi di questo genere sono regolarmente attivi dal 1993 tra le varie sedi dell'Ateneo.

Caso 3: è la versione multimediale della biblioteca, cioè una ***mediateca*** in cui sono consultabili prodotti multimediali (in genere offline) a supporto dei contenuti indicati dal docente.

Caso 4: è la didattica basata su ***videocassette*** (più recentemente su ***CD-rom***), in cui è possibile vedere o rivedere il docente che fa la sua lezione ripreso da una telecamera (in Italia uno dei casi più noti riguarda il Consorzio Nettuno [8]); è anche il principale modello per la formazione aziendale classica, privilegiando il lavoro del singolo e l'approccio detto ***CBT*** (Computer Based Training).

Caso 5: è la situazione del ***gruppo di studio***, più o meno attrezzato sul piano multimediale e delle connessioni web.

Caso 6: è la tipica situazione del ***workshop*** o della videoconferenza aziendale; i vari partecipanti possono/debbono condividere alcuni strumenti; è quindi necessaria una "piattaforma" che consenta la condivisione.

Caso 7: non c'è un modello rilevante per questa situazione (come sempre, uno schema forza un po' le cose ...); quello che forse si avvicina di più è il ***forum***, con lo scambio asincrono di messaggi.

Caso 8: è il caso che ci interessa maggiormente, l'***on-line learning***; in esso i soggetti coinvolti sono, come vedremo, molti e diversi, con differenti localizzazioni e funzioni; le modalità d'interazione sono assai importanti e di conseguenza è fondamentale la scelta della strumentazione (in particolare della piattaforma), che deve essere fortemente legata alle esigenze che si vogliono supportare.

I casi 1/2/4/8 sono i più interessanti, perchè testimoniano l'evoluzione temporale e gli sforzi fatti per inserire i nuovi strumenti (la TV, il computer, le reti, la multimedialità) nei processi didattici, a partire dagli anni Sessanta fino ad oggi. Su questi argomenti e sulle tecnologie didattiche in generale si può consultare il sito dell'ITD-CNR [11].
Lo scopo dello schema di figura 1 è di mettere in luce il fatto che quando si parla di nuove tecnologie per la didattica si possono intendere progetti e situazioni tra loro molto diversi: non ha quindi senso esprimere posizioni generali, di appoggio o di dissenso, senza riferirle ad un preciso contesto.
In sintesi direi che la principale caratteristica dei modelli (e sono più d'uno) di web learning sta nell'articolazione degli attori, più che nelle possibilità tecnologiche: in particolare la figura del tutor appare cruciale, forse più ancora di quella del docente.

## 3. QUALCHE REGOLA PER ORIENTARSI

Le considerazioni che seguono sono derivate dalla mia esperienza nella gestione del progetto "Laurea on-line del Politecnico" (LLP), sviluppata in collaborazione con Somedia (società del gruppo Espresso-Repubblica) e le cui caratteristiche principali sono descritte in [7], ma anche dalle varie attività del Centro METID a supporto dei docenti dell'Ateneo nei loro progetti di innovazione didattica[1].
Mi riferisco qui a processi didattici di tipo universitario e post-universitario (per esempio master) più che alla formazione aziendale: quest'ultima mi appare prevalentemente individuale e mirata all'ottenimento di "informazioni" per la gestione di situazioni, mentre la prima mi sembra più legata all'acquisizione di "modelli mentali" attraverso percorsi di apprendimento collettivo (la classe). Naturalmente un'integrazione tra i due modelli sarebbe utile a entrambi: offrirebbe alla formazione universitaria una robusta iniezione di pragmatismo e di "customer care", a quella aziendale una solida intelaiatura di obiettivi generali e di tappe parziali.
In ciò che segue è bene tenere a mente una distinzione tra didattica completamente on-line e didattica "mista" (in presenza e on-line). Nel primo caso c'è da aggiungere agli altri il problema della valutazione dello studente e della certificazione dei contributi da lui inviati, mentre nel caso misto ciò appare secondario.
Per le consuete esigenze di sintesi indicherò ora quelli che mi sembrano i nodi principali da sciogliere (nel caso dell'on-line learning), dedicando a ciascuno di essi solo poche righe.
❑ ***Responsabilità***. È importante definire una responsabilità precisa e univoca. Di fronte al sorgere delle ormai moltissime "agenzie" che promettono formazione con tempi brevissimi e modi facili, la questione centrale è quella di far capire che c'è chi governa il processo, assumendosi nei confronti degli studenti la responsabilità dei percorsi formativi, del controllo dell'apprendimento, della valutazione dei risultati. Un processo didattico non può mai essere "facile" (e non deve prometterlo). Esso deve dotarsi di un insieme di organismi che, con differenti compiti e livelli di responsabilità, rispondano del progetto e delle scelte didattiche. Inoltre, per un progetto di questo tipo ci vuole una notevole dose di entusiasmo e di gusto per l'innovazione, da quella didattica a quella tecnologica, da quella gestionale a quella relazionale: serve quindi un "motore". Il risultato dipende (anche) da questo.

[1] In particolare: lezioni in teledidattica tra le 7 sedi lombarde del Politecnico, produzione di CD-rom, messa on-line di materiali, web-conferences, un portale per il supporto alla didattica in presenza, un sistema per la fruizione delle lezioni da parte dei disabili, erogazione di master e corsi brevi, collaborazioni italiane ed europee.

## INGEGNERIA INFORMATICA ON-LINE AL POLITENICO DI MILANO

### Il primo corso di Laurea On-line in Italia

La Laurea in Ingegneria Informatica On-line è un percorso di studio gestito completamente a distanza attraverso il supporto di Internet. Il corso è perfettamente equivalente a quello tradizionale, con le stesse materie di studio, gli stessi crediti formativi, gli stessi carichi di lavoro, gli stessi esami e riconoscimenti della laurea "in presenza" ed è tenuto anch'esso da docenti del Politecnico di Milano. In più, il formato didattico on-line offre agli studenti la possibilità di:

- specializzarsi nell'ICT (Information Communication Technology), utilizzando per lo studio strumenti e mezzi dell'ICT stessa
- personalizzare tempi e modalità di studio senza i vincoli della presenza in aula
- scegliere un percorso di studi su misura a seconda del carico di lavoro che si è in grado di sostenere
- comunicare direttamente via Internet e via mail con tutor e docenti
- socializzare con facilità ed elevata frequenza con gli altri studenti della classe virtuale di appartenenza.

### Piano di studi standard (60 crediti/anno)

Gli studenti devono sostenere circa 8 esami all'anno (4 per semestre), per una totalità di 27 esami in 3 anni, al termine dei quali potranno scegliere se continuare gli studi e frequentare i 2 anni di "specializzazione" oppure se fermarsi alla laurea di primo livello.

| 1° ANNO |
|---|
| **Titolo dell'insegnamento** |
| Elementi di Analisi matematica (A) e di Geometria |
| Fisica sperimentale 1 |
| Fondamenti di Informatica 1 |
| Chimica A |
| Analisi matematica B (per il settore dell'informazione) |
| Fondamenti di Informatica 2 |
| Fisica sperimentale 2 |
| Economia e organizzazione aziendale B |
| **2° ANNO** |
| Elettrotecnica A |
| Fondamenti di automatica (per il settore dell'informazione) |
| Fondamenti di telecomunicazioni |
| Prova di lingua straniera |
| Fondamenti di elettronica |
| Calcolo delle Probabilità |
| Fondamenti di ricerca operativa B |
| Ingegneria del software |
| Attività integrative di laboratorio o progetto |

| 3° ANNO |
|---|
| Automazione industriale |
| Calcolatori elettronici |
| Gestione aziendale B |
| Basi di dati 1 |
| Insegnamento a scelta (area Internet) ([3]) |
| Attività integrative di laboratorio o progetto ([2]) |
| Ingegneria e tecnologia dei sistemi di controllo |
| Insegnamento a scelta (area telecomunicazioni) ([3]) |
| Insegnamento a scelta ([3]) |
| Tirocinio e prova finale |
| ([2]) Le attività integrative di laboratorio o progetto sono associate a un Corso, su proposta dello studente. |
| ([3]) I corsi tra cui lo studente potrà scegliere saranno definiti successivamente. |
| **Totale crediti nei 3 anni: 180** |

### Come funziona

Il Politecnico di Milano, istituzione universitaria tra le più prestigiose in Italia e da sempre all'avanguardia per contenuti e metodi di insegnamento, e Somedia, società privata da anni impegnata nell'organizzazione e gestione di corsi di formazione, tradizionale e on-line, hanno realizzato con la Laurea On-line un formato didattico che integra in modo coerente tutti gli aspetti chiave di un'esperienza di formazione on-line:

- didattica
- tecnologia
- organizzazione.

Allo studente della Laurea On-line viene proposto di partecipare alle attività di una classe virtuale di circa 20-25 studenti seguita da un docente e da un tutor per ciascuna materia. Le principali attività proposte alla classe sono:

**1.** studio individuale (sviluppato secondo il ritmo suggerito dai docenti) basato su:
CD-ROM che propongono i contenuti di base di ciascun insegnamento in versione multimediale
materiali on-line:
SAI (Software di Autointerrogazione)
esercizi commentati
risorse specifiche messe on-line dai docenti durante il corso (slide, spiegazioni, appunti)

**2.** *forme di apprendimento collaborativo on-line*, insieme ad altri studenti, tutor e docenti utilizzando:
strumenti asincroni (forum, e-mail, ecc.)
strumenti sincroni (sessioni live)

**3.** *prove in itinere on-line* periodiche su ciascuna materia.
Alla fine di ogni semestre lo studente dovrà sostenere gli esami in presenza presso la sede di Como del Politecnico di Milano.

❑ ***Progettazione continua***. La formazione on-line coinvolge una molteplicità di soggetti (docenti, tutor, responsabili tecnici, editor multimediali, gestori di rete, ecc.), sul coordinamento dei quali si basa il successo dell'operazione: è necessario raggiungere un certo grado di omogeneità nell'offerta didattica, pur salvaguardando la specificità dei contenuti. La modalità di erogazione (cioè tempi, verifiche, monitoraggio) è un altro elemento chiave. È utile, e spesso necessario, dedicare una struttura specifica a questo scopo, con il compito preciso di governare la complessità del progetto anche attraverso il dialogo con gli studenti. Serve un piano d'azione chiaro, ma anche modificabile in base ai feedback forniti dal sistema (cioè da studenti, docenti, tutor). L'informazione dev'essere costante e ricca (e qui si dovrebbe aprire il problema della reportistica), deve riguardare sia il gruppo che i singoli (permettendo allo studente di avere suoi momenti di autovalutazione), deve essere analizzata verticalmente (su un certo insegnamento) ma anche orizzontalmente (sull'andamento generale dello studente). Per usare una definizione cara agli specialisti dei sistemi di controllo, si tratta di un sistema di gestione "in anello chiuso", in cui l'attività viene corretta secondo i feedback che arrivano. Questa modalità è efficace solo se dispone di continue informazioni sullo stato del sistema: è però la sola che consenta un reale controllo.

❑ ***Apprendimento collaborativo***. È un aspetto cruciale, che distingue il CBT dal web learning. Si tratta di esaltare i momenti di interazione dello studente con i docenti, con i tutor, ma anche con gli altri studenti: ciò che nei videogiochi è la ricerca dell'interattività diventa qui la necessità dell'interazione. Per questo è opportuno creare delle "classi virtuali" di 20-25 studenti, ognuna seguita da un tutor per ogni materia, con momenti di lavoro collettivo sia di tipo asincrono (progetti comuni) che di tipo sincrono (sessioni "live") a cadenza predefinita. Gli studenti di un progetto di web learning sono, in genere, molto motivati e disposti a collaborare: dopo un po' è probabile che ci sia uno sviluppo di queste comunità virtuali rispetto all'obiettivo comune di aiutarsi (non a copiare, ma a risolvere i mille frangenti organizzativi e tecnologici che si presentano ogni giorno). Allo scopo di individuare tempestivamente e seguire gli studenti in difficoltà può essere utile l'attivazione di una figura di "tutor generale".

❑ ***Ritmo*** (***flessibile ma fissato***). Uno dei vantaggi della formazione on-line rispetto a quella in presenza è la quasi totale mancanza di vincoli logistici (orari e aule): ciò permette di proporre delle "finestre" nelle quali ciascun utente può collocare i suoi momenti di fruizione. Ma deve essere, nuovamente, un processo controllato. Per esempio, nel caso della LLP viene messa in linea una "agenda" settimanale in cui i docenti segnalano le parti da studiare e inseriscono gli appuntamenti. La flessibilità consente poi di definire percorsi dedicati: con il progetto LLP il Politecnico ha inaugurato alcuni piani di studio part-time, pensati proprio per quegli studenti (e sono la maggioranza) che non vogliano abbandonare i loro impegni di lavoro.

❑ ***Materiali***. È un discorso cruciale. Non basta scaricare dalla rete materiali già esistenti, è necessario riprogettarli: normalmente, infatti, i materiali didattici disponibili in rete sono utili come supporto del docente in aula, quindi come materiali integrativi. Se invece parliamo di un percorso interamente on-line è necessaria la definizione di formati e tempi di erogazione autonomi. Il docente definisce lo story-board, ma è l'editor multimediale la figura su cui grava il compito di rendere i contenuti fruibili: ruolo non esecutivo ma realmente progettuale, che richiede di operare a stretto contatto con il docente interpretandone le intenzioni e prevedendone gli utilizzi. Inoltre, è utile fornire input differenti: si tratta di sfruttare tutto il potenziale della multimedialità, dalla valorizzazione della rete (prevedendo percorsi guidati dal docente) alla integrazione di vari formati didattici (on e offline). Nel caso della LLP gli input allo studente seguono tre strade diverse, con reciproci "rinforzi": dei CD-rom inviati preventivamente, dei materiali messi in rete periodicamente, delle sessioni "live" settimanali in cui la classe intera discute con il docente o il tutor. I continui rimandi sono funzionali all'idea, ormai largamente accettata, che modalità diverse di esprimere lo stesso contenuto siano più funzionali all'apprendimento.

❑ ***Approccio***. Il "learning by doing" viene spesso citato nei progetti di web learning. Su questo punto ci sono però le difficoltà mag-

giori, per il fatto che i contenuti sono forniti da docenti che (con qualche eccezione) considerano più importante l'insegnamento dell'apprendimento: l'attenzione è quindi sulle cose che il docente ritiene importanti più che sul come esse vengono assimilate dallo studente. Gli sforzi fatti iniziano a produrre i loro frutti, mutuando dalla formazione aziendale modelli didattici basati sull'utente. Il pericolo però è quello di proporre grandi studi di casi e sistemi "fai da te", mentre il problema è, al solito, il controllo del processo.

❑ ***Monitoraggio***. Il controllo sulla qualità di un progetto di formazione on-line pone la questione di come valutare un servizio di questo tipo. La LLP ha scelto come organismo di valutazione un ente esterno[2]. La valutazione può essere fatta da vari punti di vista: l'efficacia didattica, quella tecnica, quella comunicativa (mentre solitamente nella fase iniziale di un progetto, nella quale prevalgono gli investimenti, è più difficile valutare l'efficienza). In questo senso, la collaborazione con un partner esterno come Somedia (che pure ha lasciato totale autonomia al Politecnico nelle scelte didattiche) ha offerto un elemento di stimolo su due fronti: quello dell'analisi delle esigenze dell'utenza e quello dell'attenzione a elementi di efficienza (dichiarando obiettivi, tempi, modalità) di solito non presenti nella formazione universitaria.

Ho detto all'inizio che è opportuno farsi guidare dalle esigenze degli utenti e non dalle tecnologie. Ecco qualche altro esempio di come alcune delle caratteristiche sopra indicate possano essere inserite all'interno di progetti di web learning (mi riferisco naturalmente a progetti che conosco bene, cioè a progetti METID).
Il sistema Corsi on-line [5], offerto ai docenti del Politecnico come supporto alla loro didattica in presenza, è un esempio della necessità di fornire, oltre che materiali in diversi formati, anche, ma dovrei dire soprattutto, servizi ai docenti: per la gestione della classe, per l'aggiornamento dei contenuti, per la comunicazione verso gli studenti, per il testing della loro preparazione. La figura 2 mostra una delle pagine


**FIGURA 2**
*Una pagina del sito Corsi on-line di METID*

del sistema Corsi on-line: nella barra di sinistra è presentata l'organizzazione dei servizi.
Il progetto FormAmbiente [9] ha l'obiettivo di addestrare funzionari regionali all'utilizzo di pacchetti software su tematiche ambientali. È organizzato con materiale in rete e momenti in presenza (corsi brevi della durata di qualche giorno). Qui l'aspetto importante è dato dalla possibilità di definire diversi "profili di utente", ciascuno dei quali ha differenti materiali e servizi di cui può usufruire. Ciò consente sia il controllo degli accessi che la differenziazione dei percorsi tra gli utenti (il discorso dei percorsi differenziati si collega a una delle critiche di Stoll su cui tornerò nelle conclusioni).
In questi esempi, come si attua l'interazione? Nel caso della LLP, essa avviene con le sessioni "live", i test on-line e le prove in itinere, la web-conference (con la possibilità di porre domande in tempo reale), l'uso delle mail, gli scambi di elaborati tra studenti e tutor, i progetti di corso (a volte fatti in gruppo); inoltre nei CD-rom sono previsti momenti di autovalutazione (le cosiddette "piazzole di sosta") e link al sito, per scaricare i risultati. Nel progetto Corsi on-line, il sistema fornisce soprattutto uno stimolo all'interazione diretta in classe (lo strumento è un supporto alla didattica in presenza). Il progetto FormAmbiente prevede un'interazione utenti-tutor da fare on-line ma anche da sviluppare nei momenti in presenza, durante i corsi brevi.
Concludo ritornando alla questione dei materiali. Non servono la spettacolarizzazione e gli effetti speciali; serve invece un'attenta taratura dei contenuti e dei tempi di erogazione, in relazione agli utenti e alle loro esigenze. Un sito altamente multimediale può ri-

[2] Si tratta dell'Osservatorio sulla Comunicazione dell'Università Cattolica di Milano.

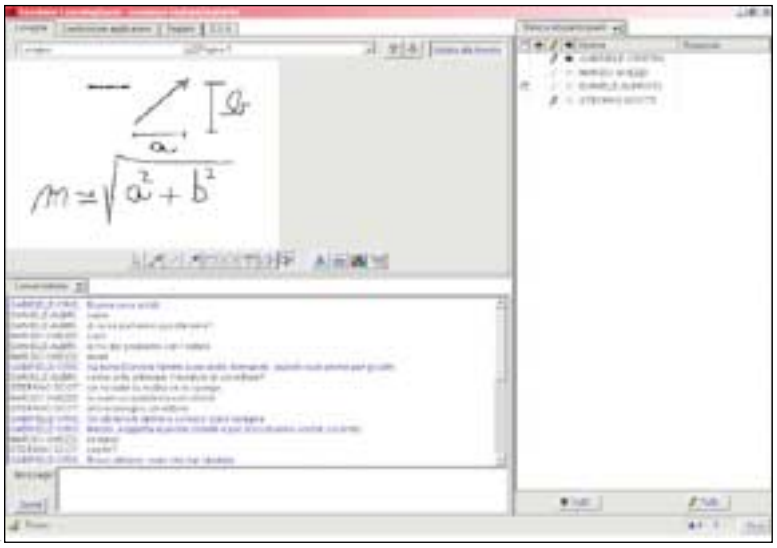
chiedere elevati tempi di accesso e di fruizione, magari incompatibili con un percorso formativo lungo, mirato a utenze che possono avere connessioni e macchine differenti. Meglio allora una scelta che veicola alcuni contenuti (quelli più assestati) attraverso CD-rom e altri (quelli maggiormente dinamici) via rete.

## 4. A COSA CI SERVE LA TECNOLOGIA (E QUANTO CI COSTA)

Che il computer di ciascuno degli attori di un processo di web learning sia equipaggiato con dei software "user friendly" lo do per scontato. Il punto è un altro: come direbbe Norman, si tratta di "rendere più vicino il giorno in cui la tecnologia informatica scomparirà dalla vista e la nuova tecnologia che ne prenderà il posto sarà immediatamente assimilabile e facile da usare ..." [13]. In attesa di questo momento (ma Stoll dubita che ci sarà mai), si tratta di piegare la tecnologia alle esigenze degli utenti. A questo punto il nostro obiettivo è abbastanza facile da enunciare: la tecnologia deve fornire gli strumenti (dalla banda larga alla tavoletta grafica) per favorire l'interazione.
Procedendo anche qui in maniera un po' schematica, ecco le principali ***necessità***:

- interazione uno-a-uno (studente-docente);
- interazione di gruppo (tra gli studenti);
- accesso alle informazioni e scaricamento dei materiali;
- monitoraggio e tracciamento (del singolo e della classe).

**FIGURA 3**
*La piattaforma Learning Space durante una sessione "live"*



A queste esigenze corrispondono alcuni strumenti ben noti dell'on-line learning, più o meno presenti in tutte le piattaforme [3]. Essi sono:

- la ***condivisione della lavagna*** (i partecipanti alla sessione di lavoro devono poter intervenire reciprocamente; è presente un livello superiore di gestione);
- la ***chat***, sia testuale che audio (almeno in broadcast), mentre è meno importante quella video;
- la ***bacheca***, con una duplice funzione (affissione di messaggi in uscita e deposito di testi in arrivo);
- il ***forum***, come luogo di dibattito "asincrono" (sia un forum generale che un forum per ogni materia);
- un sistema di ***test in tempo reale*** (con la possibilità di vedere immediatamente risultati e statistiche);
- un sistema di ***mail dedicate*** al progetto (indipendenti dalle mail personali degli utenti);
- la ***web-conference***, possibilmente integrata con il sistema di messaggistica della piattaforma (al fine di permettere domande e risposte in tempo reale);
- la definizione di differenti "***profili di utenza***", con la possibilità di collegarli a un database nel quale convergano informazioni quali-quantitative sul singolo utente;
- un sistema semplice e rapido di ***reportistica***, sul singolo, sulla specifica prova, sulla classe (è in pratica una matrice tridimensionale).

Altri servizi, come quello di aggiornamento rapido della propria pagina web o tutti quelli di gestione della classe, possono ulteriormente favorire il processo: ma una piattaforma che fornisca in modo affidabile i servizi sopra elencati può bastare.
In figura 3 è mostrata l'interfaccia della piattaforma Learning Space utilizzata nelle sessioni "live" della LLP: sono visibili alcuni degli strumenti indicati (la lavagna condivisa, la chat, il test in tempo reale, la gestione degli studenti presenti).
Quali sono i costi di un progetto di web learning ? Per rispondere a questa domanda è necessario distinguere tra costi di investimento (in hardware e software), costi di produzione dei materiali, costi di erogazione. Esistono poi, essenzialmente per i docenti, dei costi non monetizzabili.
Il dettaglio dei costi monetari dipende fortemente dalla dimensione, dal target di utenza e dalla copertura che si vuole dare al proget-

to. Come termine di paragone si può prendere la Cardean University americana [4], creata da un pool di università e di imprese dell'area ICT e diretta da D. Norman: nella sua realizzazione sono stati investiti circa 50 milioni di euro. Il progetto LLP ha richiesto un investimento di circa 5 milioni di euro (in larga misura sostenuti da Somedia, partner del Politecnico). I costi nella fase di avvio sono elevati: il recupero degli investimenti non può, quindi, che richiedere tempi lunghi e un mercato in crescita.
Anche per quanto riguarda i tempi di produzione la variabilità è elevata, tuttavia qualche indicazione può essere fornita. Il rapporto fruizione/preparazione va da 1:30 a 1:100 (per un'ora di fruizione possono essere necessarie da 30 a 100 ore di preparazione), con una forte dipendenza dalla modalità/rigidità del processo produttivo adottato. In ogni caso, per il docente il rapporto non è mai meno di 1:10. Progetti didattici del tipo videocassette (il caso 4 di figura 1) sono più veloci: per esempio, il Consorzio Nettuno dichiara un rapporto tra 1:20 e 1:30.
Un tempo di preparazione di questa rilevanza richiede ovviamente un team formato dal docente e da un gruppo di collaboratori: nei vari casi a cui ho lavorato il team era formato in media da 3-5 persone, con compiti diversificati (registrazione dell'audio, costruzione del glossario, analisi dei siti interessanti, creazione degli esercizi, ecc.).
Per i vari motivi indicati, lo sforzo organizzativo non è mai modesto. La gestione del processo produttivo è quindi un punto delicato. Per il team dei docenti contano assai di più le motivazioni didattiche e le condizioni di lavoro che non la remunerazione economica: qualunque intervento didattico "in presenza" avrebbe un rapporto tra ore di preparazione e retribuzione molto più favorevole.
Un discorso a parte va fatto per i costi non monetizzabili: ne accenno qui utilizzando un modello della Teoria dei giochi noto come "Il dilemma del prigioniero", presentato per esempio in [2], che in questo contesto si potrebbe ribattezzare "Il dilemma del (giovane) docente".
Lo schema del modello riporta una matrice dei benefici di due giocatori, ciascuno dei quali ha davanti a se la scelta (da ripetere ogni giorno) tra due alternative: cooperare con il suo antagonista o defezionare, cioè fare i propri escusivi interessi. I due non possono comunicare tra loro. Il sistema giudiziario premia il pentitismo, cioè la defezione.
Se i due cooperano i benefici per entrambi sono misurabili in 3 unità (in una scala tra 0 e 5). Se uno dei due sceglie un atteggiamento cooperativo e l'altro no, colui che defeziona ottiene il massimo beneficio (5) mentre l'altro viene duramente punito (beneficio 0). Se infine entrambi defezionano i benefici sono molto minori (il risultato è di 1 per entrambi): ma si tratta dell'unico punto di equilibrio. La strategia più logica per ognuno dei due è quindi quella di defezionare (questa scelta domina la scelta di cooperare), ma la situazione di mutua cooperazione sarebbe più favorevole per entrambi: se i due si fidano l'uno dell'altro la attueranno, se no avranno un beneficio minore (1 invece di 3).
In figura 4 e nel caso che voglio descrivere, i giocatori sono due (giovani) docenti, il sistema è quello della promozione universitaria, cooperare tra loro vorrebbe dire che entrambi adottano una didattica innovativa, defezionare vuol dire muoversi secondo le regole del sistema universitario (che premia quasi solo la ricerca) cioè minimizzare l'impegno didattico attuando una didattica tradizionale. La coppia di scelte dei due giocatori produce uno dei quattro risultati indicati nella parte (a) di figura 4.
Cosa c'entra tutto ciò con i costi di un progetto di web learning ? E perchè il modello vale per un "giovane" docente ? Nel nostro caso, cooperare significa "spendersi" nel progetto di web learning, defezionare significa adottare una didattica classica (meno dispendiosa in termini di tempo). Se il suo vicino di stanza non fa altrettanto, al giovane docente non conviene cooperare perchè impegnandosi in questa direzione scriverà meno articoli scientifici e farà meno carriera (se è giovane questo conta, se è già arrivato all'apice il modello ha meno significato). Se però entrambi decidessero che vale la pena di lavorare a un simile progetto, i benefici per la didattica di entrambi aumenterebbero e nel contempo non ci sarebbero danni in termini di carriera. Se poi qualcuno per loro (cioè sopra di loro) decidesse per la cooperazione reciproca il successo sarebbe garantito. Questo è il ruolo che l'istituzione universitaria può avere per far decollare progetti di web learning: modificare il quadro (passando

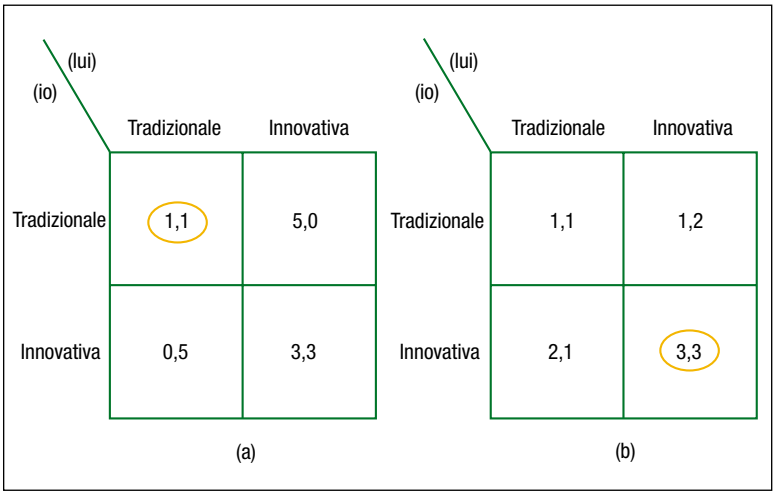



**FIGURA 4**
*La matrice di payoff per il dilemma del giovane docente*

cioè a una situazione come quella descritta nella parte (b) della figura), in modo da garantire visibilità e benefici immediati ai progetti di didattica innovativa.

## 5. EURO E NON SOLO

Cosa c'è in giro per il mondo nel settore di cui parliamo ? Una doverosa premessa è che qualsiasi esame è destinato ad una rapida obsolescenza, a causa della dinamica con cui le cose si stanno muovendo in questo momento. In generale si possono individuare tre tipologie di erogazione [15, 19]:

- università a distanza, cioè istituzioni accademiche che offrono esclusivamente corsi a distanza, sia attraverso la rete telematica (on-line learning), sia seguendo le modalità più tradizionali della formazione per corrispondenza (FaD tradizionale);
- università "dual-mode", cioè università che offrono corsi sia in presenza che on-line, in modo da cercare di soddisfare le esigenze di due fasce di utenza distinte;
- università in presenza con corsi "misti", cioè università che offrono esclusivamente corsi in presenza, alcuni dei quali prevedono una parte del programma su web (la rete ha la funzione di integrare i contenuti trattati in aula).

Cominciamo analizzando la situazione italiana. Viene spesso ricordato il Nettuno [8], un Consorzio che comprende più di 30 università, la Rai ed altri partner, da anni impegnato in un progetto didattico basato sulla produzione di videocassette (attualmente ha un catalogo che comprende molte migliaia di ore di videolezioni e di esercitazioni) diffuse attraverso due canali satellitari della Rai. Il progetto Nettuno è un ottimo esempio per il caso 4 dello schema di figura 1. I meriti storici del Consorzio non possono che essere riconosciuti. Il modello didattico, invece, mi lascia perplesso a causa di alcune differenze non marginali rispetto a quanto ho descritto in queste pagine: la poca interazione (anche se ultimamente il Consorzio ha modificato il suo modello, spostandolo in parte su Internet) e una responsabilità abbastanza diluita (quasi sempre gli esaminatori degli studenti sono docenti diversi da quelli che hanno registrato le videocassette, spesso ne ignorano il contenuto) mi sembrano i limiti più grossi dell'approccio.

Altre esperienze on-line sono presenti sul fronte dei corsi di master. Il Politecnico di Milano ha avviato alla fine del 2001 il master [16] NBA *Net Business Administration*, in collaborazione con Sfera società del gruppo Enel. La società Profingest, collegata all'università di Bologna, offre un master on-line [17] anch'esso nel settore della business administration.

Un'iniziativa particolare fa capo al Consorzio ICoN [10], costituito da un gruppo di università allo scopo di diffondere la cultura italiana all'estero: il Consorzio sta mettendo in rete varie unità didattiche e si propone di gestire attraverso le ambasciate e i consolati un sistema di certificazione per utenti residenti in ogni parte del mondo.

Ci sono poi alcuni progetti speciali che stanno avviandosi con il concorso di gruppi misti di università e imprese: tra questi cito il Consorzio LabCon, promosso dalla CRUI (la Conferenza dei Rettori), che vede una decina di atenei affiancare una società del gruppo Telecom. Un'indagine esplorativa promossa dai Presidi delle Facoltà di Ingegneria, allo scopo di creare un network di università, ha consentito di censire la situazione di quelle che già operano nel settore dell'e-learning.

Esiste poi il mercato "consumer", interpretabile solo in termini molto aggregati: l'annuale rapporto [1] dell'ANEE (l'associazione che riunisce i principali editori multimediali italiani) fornisce qualche dato.

In Europa fanno testo i numerosi documenti che la UE dedica alla formazione mediante le tecnologie dell'ICT. Le università europee di maggiore tradizione in questo campo sono la

| | Tipologia dei corsi erogati | Paesi guida | Università |
|---|---|---|---|
| Università virtuali | On-line(oppure per corrispondenza) | Canada<br>USA<br>Spagna | http://www.athabascau.ca/ (Athabasca)<br>http://www.cardean.edu/ (Cardean)<br>http://www.uoc.es (Un. Oberta) |
| Università "dual mode" | Sia in presenza che on-line | Canada<br>Australia | http://www.open.uoguelph.ca/ (Guelph)<br>http://www.csu.edu.au/ (Charles Sturt) |

**TABELLA 1**
*Alcune esperienze straniere*

Open University inglese [14] e la UOC - Universitat Oberta de Catalunya [20]. In entrambi i casi si tratta di università "general purpose", che offrono corsi su uno spettro molto ampio a studenti di varia provenienza e background. Ci sono poi vari tentativi di mettere in rete atenei di diversi stati, favorendo l'integrazione: tra i progetti in cui il Centro METID è stato direttamente coinvolto cito Teleregion-SUN (tra le regioni autodefinitesi "i motori d'Europa"), Vimims (con la partecipazione di alcune università non-UE), Corel (che ci vede collaborare proprio con la UOC), Tuelip (una rete di università con forti esperienze nell'on-line, con il supporto di IBM). In tutti questi casi si tratta di mettere in rete materiali didattici fruibili da studenti di altre università, a volte con forme di scambio. Ed è proprio il principio di reciprocità che sta alla base di questi progetti che ne è anche il punto debole: gruppi abbastanza eterogenei di atenei, creati per ottenere finanziamenti UE, difficilmente riescono a garantire le caratteristiche di qualità necessarie in un progetto di e-learning. Da qui una certa difficoltà a far decollare queste operazioni, dotandole di adeguati obiettivi didattici.
Non è quindi facile fornire le linee generali di sviluppo dei progetti europei e mondiali. Il settore è certamente in fase di espansione, anche se non così brillante come qualche previsione aveva indicato solo un anno fa. Nel mondo ci sono comunque tre situazioni degne di nota: quella americana, quella canadese, quella australiana. In tabella 1 sono riportati alcuni indirizzi significativi, riferiti alla classificazione proposta all'inizio di questa sezione.

## 6. PENTIRSI?

Le conclusioni di un articolo pongono sempre il problema di riassumere in poche righe un ragionamento articolato [6]. Provo a farlo partendo da queste premesse:
- parliamo di istruzione universitaria o post-universitaria, in cui prevale il trasferimento di informazioni (piuttosto che l'attenzione agli aspetti educativi) unito alla fornitura di metodi e strumenti per affrontare la complessità del lavoro;
- parliamo di un percorso formativo (e non di un più o meno gradevole fai-da-te), con responsabili precisi, tappe intermedie, servizi da utilizzare, valutazione e monitoraggio, momenti di confronto e di dibattito, un preciso obiettivo finale;
- parliamo a persone motivate (da ragioni varie, di carriera, di status, di curiosità, ma in ogni caso non "precettate dai genitori") per le quali è o dovrebbe essere chiaro che il percorso non è facile né divertente.

Riparto dalle principali obiezioni alla formazione on-line, in particolare da alcune fatte da Stoll: per ciascuna di esse provo a indicare i modi perchè "la goccia di pioggia finisca nel bacino giusto".

1. ***Assenza di una dimensione sociale***. Le scuole per corrispondenza e la vecchia FaD hanno fallito perchè non hanno considerato le relazioni tra i soggetti (lasciando lo studente isolato) e perchè hanno concentrato lo sforzo sui materiali (senza fornire a fianco i servizi). Bisogna invece pensare e progettare un apprendimento sia interattivo che collaborativo: spero di aver esemplificato a sufficienza quest'esigenza nell'articolo. La tecnologia lo consente, è necessario progettare la cosa fin dall'inizio.
2. ***Prevalenza del virtuale e passività***. Questa obiezione ("meglio una camminata nel bosco che la simulazione a video della crescita di una pianta") è interessante, ma riguarda tutta l'università (e non solo la FaDoL) in quanto basata molto sul sapere e molto meno sul saper fare. È vero che la formazione on-line ha in se

alcuni "germi dannosi" dell'istruzione programmata[3], ma ciò avviene prevalentemente secondo alcune modalità (per esempio quella basata sui "learning objects") che ipotizzano una parcellizzazione del sapere da riassemblare secondo le circostanze, senza una guida responsabile del percorso. Un progetto fortemente basato sull'interattività può scongiurare il pericolo della passività: la lentezza dei collegamenti è un problema in via di risoluzione e la possibilità di graduare (on/off line, sincrono/asincrono) va sfruttata.

**3.** ***Percorsi facili e superficialità***. Questa obiezione è assolutamente fondata, ma solo per qualche caso "da pubblicità". Non si dovrebbe promettere una formazione facile, da ottenere nei ritagli di tempo, ma un percorso duro, controllato, garantito nel risultato. Eventualmente si possono attivare percorsi diversi, con tempi differenziati (part-time), ma sempre sotto una precisa e unica responsabilità. Studiare non è "divertente" e la formazione on-line non è l'edutainment.

**4.** ***Valutazione e standardizzazione***. La considerazione di chi dice che nella FaD si promuove la capacità dello studente di rispondere solo ai test e alle domande standard non è priva di fondamento. In realtà è una questione di tempo e di progetto didattico, non una scelta obbligata: se il docente è disponibile a correggere elaborati più articolati (pur di superare lo scoglio della scrittura delle formule) la questione non si pone. È comunque utile che nel "registro" del docente ci sia spazio per valutazioni anche qualitative, basate sul livello della partecipazione dello studente ai forum e sul tipo di richieste che gli fa pervenire. Inoltre la valutazione finale dovrebbe essere in presenza (nel caso della LLP è così), con la doppia funzione di garantire dell'identità di chi ha inviato gli elaborati durante il semestre e di consentire un confronto faccia-a-faccia con lo studente.

**5.** ***Aspetti commerciali***. La FaDoL ha messo in moto un giro di affari crescente (anche se forse non del livello previsto), in cui si sono inserite anche le università: ciò è vero e dipende dalle molte (troppe) istituzioni universitarie costrette a competere sull'unica risorsa veramente scarsa, gli studenti. In aggiunta a ciò, la formazione on-line toglie anche il "filtro territoriale" (l'università della propria città o regione) che forniva un livello garantito di domanda. La competizione comporta dei costi e una visione attenta anche agli aspetti commerciali. Quando una cosa è gratuita (e la maggior parte delle cose che si trovano in Internet lo sono) il suo valore economico è basso: il costo elevato della FaDoL può essere anche una garanzia della sua qualità (molte volte nel progetto LLP ci siamo detti "gli studenti pagano salato e quindi hanno diritto a ...").

**6.** ***Investimenti vs. risultati***. La formazione on-line ha costi elevati, non c'è dubbio[4]. Ma i costi erano elevati anche quando si costruivano i primi edifici scolastici invece di pagare un precettore a ogni alunno (e l'obsolescenza dei contenuti non sarà così forte come si teme). Inoltre la ricerca, perchè di questo si tratta, ha alcuni elementi di rischio: perchè ciò che non si rimprovera al CERN dovrebbe essere messo in conto alla FaDoL? Il punto vero è capire a quali condizioni e come si può fare della buona formazione on-line.

Un discorso più difficile è quello sugli sviluppi futuri: le previsioni di espansione a ritmo vertiginoso sono finite insieme a tutte le altre sulle sorti del mondo web. Quello della formazione on-line è però un settore che risponde a esigenze vere: un certo ottimismo non è solo rituale. C'è poi tutto il discorso sulle tecnologie e sulla integrazione.

Recentemente si parla molto di "convergenza", cioè dell'integrazione TV-computer, e dei futuri benefici che essa produrrà. Più che le visioni un po' inquietanti (il tizio che fa home-banking mentre guarda una soap-opera e attende una pizza ordinata on-line), mi sembra plausibile e interessante lo scenario che vede l'invio dei materiali su canali a larga banda (via cavo o via etere) e i ritorni verso il docente (feedback, test volanti, domande) via web.

---

[3] Negli anni Cinquanta B.F. Skinner fu il principale promotore di una modalità di apprendimento basata su domande e successivi avanzamenti lungo un percorso programmato, una sorta di primitivo ipertesto.

[4] Stoll ([18], p.86) cita il caso di un investimento di 100 milioni di $ per un progetto che ha visto solo 100 studenti arrivare alla conclusione.

Attenzione però: oltre alla banda dovrà essere larga anche la memoria, una mediateca personale ingombra.
Come valutare dunque l'efficacia di un sistema di web learning? cioè di un insieme di contenuti culturali, materiali forniti, servizi offerti, organizzazione?
Direi essenzialmente in tre modi:
***a*** rispetto alla situazione precedente (cioè guardando alle opportunità che offre);
***b*** rispetto alle attese (cioè dal punto di vista degli obiettivi dichiarati);
***c*** rispetto agli standard (cioè attraverso un confronto con esperienze positive).
Nel nostro caso direi che l'opportunità di ri-avvicinare alla istruzione universitaria una parte di persone che avevano fatto (o stanno per fare) scelte diverse è sicuramente positiva, che l'obiettivo di un percorso facile è sbagliato ma non lo è quello di una buona interazione con docenti e tutor (certamente ottenibile e più favorevole di molte situazioni in presenza), che sono ancora in corso le definizioni di standard adeguati (e questo è un buon motivo per esserci).
Tutto bene, dunque ? No, certamente. Ho ben più di un dubbio su tutta questa vicenda. Ma d'altra parte, anche senza andare a casi eclatanti come la bomba atomica o la clonazione, non ho mai visto una scoperta o una tecnologia (o più semplicemente un'opportunità) che sia stata rifiutata per ragioni etiche o comunque concettuali. Il massimo che si può fare è volgerla in opportunità utile e cercare di dominarne gli effetti indesiderati. Meglio quindi operare criticamente fin dall'inizio piuttosto che scrivere libri di pentimento 20 anni dopo.

## Ringraziamenti

Molte persone hanno contribuito alle mie analisi. Tra le tante, ringrazio in particolare coloro che lavorano al Centro METID: le esperienze e le riflessioni che insieme abbiamo maturato sono la base di questo articolo.

## Bibliografia

[1] ANEE: *Editoria, contenuti e servizi nell'economia digitale in Italia*. Milano, 2001.

[2] Axelrod R: *Giochi di reciprocità*. Feltrinelli, 1985.

[3] Biolghini D, Cengarle M(eds.): *Net-learning*. Etas, 2000 (contiene un'analisi delle principali piattaforme).

[4] Cardean University, http://www.cardean.edu

[5] Centro Metodi E Tecnologie Innovative per la Didattica, http://corsi.metid.polimi.it, Politecnico di Milano.

[6] Colorni A: *Così nasce un "cyberateneo"*. Il Sole 24ore, 28 settembre 2001.

[7] Colorni A, Della Vigna P, Negrini R: *Www.laureaon-line.it: the on-line bachelor programme in computer engineering at Politecnico di Milano*. ICSEE 2002, 2002.

[8] Consorzio Nettuno, http://www.nettuno.stm.it/nettuno/index.htm

[9] Formez, http://www.formambiente.org, progetto di formazione su temi ambientali per i funzionari regionali.

[10] ICoN – Italian Culture on the Net, http://www.italicon.it

[11] Istituto per le Tecnologie Didattiche del CNR, http://itd.ge.cnr.it

[12] Laurea on-line in Ingegneria Informatica, http://www.laureaonline.it, Politecnico di Milano.

[13] Norman D: *Il computer invisibile*. Apogeo, 2000.

[14] Open University, http://www.opern.ac.uk

[15] Padovani N: *La formazione a distanza di livello universitario: il primo caso italiano di laurea ondine*. Tesi di laurea, relatore prof. F. Colombo, Università Cattolica del Sacro Cuore, Milano, 2001.

[16] Poliedra, Sfera, http://www.poliedra.polimi.it/l_master.htm, Net Business Administration.

[17] Profingest, Executive MBA on-line, http://www.profingest.it

[18] Stoll C: *Confessioni di un eretico high-tech*. Garzanti, 2001.

[19] Trentin G: *Insegnare e apprendere in rete*. Zanichelli, 1998.

[20] Universitat Oberta de Catalunya, http://www.uoc.es

Alberto Colorni è professore di Ricerca Operativa e direttore del Centro METID (Metodi E Tecnologie Innovative per la Didattica) al Politecnico di Milano.
Le sue ricerche vanno dai modelli matematici di decisione, alle applicazioni all'ambiente e ai trasporti, ai processi di e-learning.
È autore di circa 200 pubblicazioni; ha diretto progetti di ricerca italiani ed europei.
Ha vinto il premio Philip Morris per la Ricerca Scientifica nel 1989 e il Premio Cenacolo per l'Editoria e l'Innovazione nel 2001.
E-mail: alberto.colorni@polimi.it

# SISTEMI OPERATIVI "OPEN SOURCE": IL CASO LINUX

**Giampaolo Amadori**
**Gianfranco Bazzigaluppi**
**Walter Bernocchi**
**Mauro Gatti**

In questo articolo vengono presentate le caratteristiche di *LINUX* dalla sua storia, all'evoluzione fino all'affermarsi all'interno del più generale fenomeno del software "*Open Source*". Si dimostra come LINUX sia già pronto per il mercato e come permetta la realizzazione di soluzioni di elevata affidabilità, prestazioni, scalabilità, flessibiltà , sicurezza ed economicità. A tale riguardo si analizzano ed approfondiscono le soluzioni basate su *CLUSTER LINUX* che si stanno affermando sul mercato.

## 1. BREVE STORIA DEI SISTEMI OPERATIVI CON PARTICOLARE RIFERIMENTO ALLO UNIX

In origine il termine "software" (SW) veniva usato per identificare quelle parti di un sistema di calcolo che fossero "modificabili" liberamente tramite differenti configurazioni di cavi e spinotti, diversamente dall'"hardware" (HW) con cui si identificava la componentistica elettronica. Successivamente all'introduzione di mezzi linguistici atti ad alterare il comportamento dei computer, SW prese il significato di "espressione in linguaggio convenzionale che descrive e controlla il comportamento della macchina". In particolare occorre notare che ogni macchina aveva un suo proprio linguaggio convenzionale, comunemente detto Assembler.
Nella relativamente breve storia della "Information Technology", tutte le società produttrici di HW hanno cominciato a fornire, assieme all' HW della macchina un certo numero di programmi in grado di svolgere le funzioni di base. Inizialmente si parlava di Monitor, poi di Supervisori, infine è stato adottato il nome di "Sistema Operativo". Tali programmi erano scritti sempre in Assembler, cioè nel linguaggio specifico della singola macchina.
Una eccezione a tale regola è stata il sistema operativo UNIX, realizzato da una organizzazione che non produceva HW, e precisamente i Laboratori Bell, e che pertanto fu progettato fin dall'inizio per risultare indipendente dalla specifica piattaforma HW su cui era stato inizialmente scritto. Non solo, i progettisti iniziali dello UNIX, proprio per renderlo indipendente dalla piattaforma HW, svilupparono un linguaggio, conosciuto come "Linguaggio C", che contiene costrutti della programmazione strutturata (come i cicli FOR e DO...WHILE, le clausole IF...THEN...ELSE, ecc.) oltre a permettere la gestione precisa delle posizioni di memoria, possibile con l'Assembler.
Un'altra peculiarità, è che, avendo allora (anni '70) la Bell una causa in corso per violazione della legge statunitense sui monopoli, lo UNIX veniva inizialmente distribuito corredato dei sorgenti. Questo permise ad altri di portare il sistema UNIX stesso su piattaforme diverse da quelle usate dai laboratori

Bell e ne causò una rapidissima diffusione fra le comunità dei ricercatori e degli sviluppatori. Non solo: la disponibilità dei sorgenti permetteva a chiunque di contribuire al miglioramento ed all'espansione delle funzioni del sistema operativo. In pratica si formò una comunità di sviluppatori volontari, in genere dell'ambiente universitario, ma non solo, che contribuì, in un modo di operare che potremmo definire "collaborativo", alla crescita dello UNIX.
Nel 1984 la Bell perse la causa e la proprietà dei laboratori UNIX passò all'AT&T, che iniziò a rivendicarne i diritti, cioè cominciò a chiedere delle royalty, ma soprattutto mise un freno alla distribuzione dei sorgenti. Ne seguì una battaglia per i diritti di proprietà dello UNIX che durò una decina d'anni. Un folto gruppo di costruttori di sistemi informatici costituirono la Open Software Foundation, con l'obiettivo di farla diventare la proprietaria del software di base comune a tutti e sul quale tutti avrebbero costruito il loro sistema UNIX. Il risultato fu che quell'unico UNIX divenne una molteplicità di Sistemi Operativi proprietari: AIX, Digital UNIX, HP-UX, Sinix, Irix, UNIX386, ecc.

## 1.1. Il progetto GNU e la Free Software Foundation

La principale conseguenza della scomparsa della Bell e della trasformazione di UNIX in tanti sistemi proprietari fu che la comunità di programmatori, formatasi spontaneamente in quegli anni si trovò impossibilitata a continuare a lavorare. Uno degli "hacker" più lungimiranti, Richard M. Stallman, ricercatore presso il Laboratorio di Intelligenza Artificale dell' M.I.T, nel 1985 diede origine al progetto GNU ed alla Free Software Foundation.
Vorrei sottolinerare come l'uso recente della parola "hacker" con il significato di "pirata informatico" è una deformazione dovuta ai mass media. I veri "hacker" rifiutano questo significato e continuano ad usare la parola intendendo "Qualcuno a cui piace programmare e gode nell'essere bravo a farlo" [9].
L'acronimo GNU significa "GNU is Not Unix", ad indicare che l'evoluzione che aveva seguito il sistema operativo UNIX era sbagliata.
La Free Software Foundation (FSF) si occupava e si occupa tuttora di eliminare le restrizioni sulla copia, sulla ridistribuzione, sulla comprensione e sulla modifica dei programmi per computer, concentrandosi in particolare sullo sviluppo di nuovo software libero, inserendolo in un sistema coerente che possa eliminare il bisogno di utilizzare software proprietario.
Stallman sì premurò anche di concepire un'infrastruttura legale entro cui il sistema GNU potesse prosperare. Introdusse la **licenza GPL**, GNU Public License o General Public License, che garantisce le seguenti libertà:
- libertà di eseguire il programma per qualunque scopo, senza limitazioni;
- libertà di adattare il programma alle proprie necessità (l'accesso ai sorgenti è condizione essenziale);
- libertà di copiare il programma senza limitazioni;
- libertà di distribuire le copie modificate (l'accesso ai sorgenti è condizione essenziale).

Oltre a questi diritti, stabilisce un dovere: tutte le modifiche a software licenziato con la GPL devono essere rilasciate con la stessa licenza. La GPL è quindi una licenza persistente.
Il concetto di **Copyleft** - all rights reversed (gioco di parole con Copyright - all rights reserved) è stabilito nelle quattro libertà e nel dovere sanciti nella GPL e che vengono concessi (*left*, participio passato di leave).

## 1.2. Il kernel LINUX

La Free Software Foundation si occupò di riscrivere tutto il sistema UNIX per non dovere royalty a nessuno. Tutti i componenti venivano però innestati su un kernel UNIX, essendo il progetto del Kernel GNU (di nome Hurd) non ancora completato. A tutt'oggi il Kernel Hurd è ancora in fase di sviluppo.
Il 25 agosto 1991 uno studente finlandese, Linus Torvald, annunciò che stava lavorando ad un kernel UNIX-like ("non un progetto professionale come GNU" diceva il suo messaggio) e chiedeva suggerimenti su quali funzioni erano ritenute interessanti; e fornì quindi questo kernel, in seguito ribattezzato LINUX, alla Free Software Foundation.
Il **sistema GNU/Linux** (l'insieme del kernel Linux e delle componenti del progetto GNU) oggi è un sistema operativo completo, funziona su un numero notevolissimo di piat-

taforme hardware. Senza campagne pubblicitarie sponsorizzate da organizzazioni commerciali, ma solo per forza sua, è arrivato in pochi anni ad oltre dieci milioni di installazioni nel mondo. In particolare si stima che Linux sia il motore di oltre sei milioni di server Web in Internet.

## 1.3. Software Libero (Free software) ovvero "Open Source"

Un altro nome per il software libero è "Open Source". Questo termine è nato il 3 febbraio 1998 in una riunione a Palo Alto, California, a cui partecipavano Todd Anderson, Chris Peterson (del Foresight Institute), John "maddog" Hall and Larry Augustin (ambedue di Linux International), Sam Ockman (del Silicon Valley Linux User's Group), and Eric Raymond. Il concetto venne sviluppato durante una discussione sul doppio significato, in inglese, della parola "free": se usata nella frase "free beer tonight", significa "gratis", ma se usato come nel primo emendamento della Costituzione Americana con il concetto di "free speech right" allora significa "diritto alla libertà di parola".
Per evitare questa possibile confusione fu coniato il termine "Open Source". (La storia completa della nascita e della diffusione del termine si può trovare in [2] che, oltre a contenere la storia di OSI, intesa come Open Source Initiative, fornisce anche molti puntatori a documenti che riguardano la storia degli hacker e di Linux.)

### 1.3.1. Il Software Libero è più affidabile e più efficiente

Il vantaggio principale del software libero non consiste, come invece molti vogliono credere, nel fatto che è a basso costo perché si può copiare senza limiti e senza royalty. Il vantaggio principale sta nella sua robustezza. Questa robustezza è dovuta al modo come viene sviluppato il software, che è diverso dal modo di sviluppo del software proprietario. I moduli del software libero vengono sviluppati da poche persone, spesso da una sola persona, ma poi, invece di essere protetti e tenuti nascosti, come accade per il software proprietario, vengono pubblicati in Internet, dove sono a disposizione di tutti i programmatori che desiderano rendersi utili verificando software altrui. In pratica il software viene sottoposto a decine, centinaia di revisioni da parte di professionisti. Eventuali buchi o debolezze vengono evidenziati molto prima che possano emergere durante il funzionamento. Volendo banalizzare si tratta dell'applicazione del vecchio proverbio: "quattro occhi vedono meglio di due", che in questo caso diventa, in inglese, "Many eyes make all bugs shallow".
Per lo stesso motivo il software libero può essere anche più efficiente. Se un hacker, inteso come professionista del software, esaminando una porzione di codice si accorge che esiste un modo più efficiente di risolvere un problema, molto probabilmente lo comunicherà alla comunità degli sviluppatori, magari inviando la parte di codice già scritta. Alla lunga questo meccanismo spontaneo porta a software robusto ed efficiente.
Nel mondo del software libero inoltre il software viene rilasciato senza l'ansia di rispettare le scadenze di annuncio, tipiche del mondo del software proprietario. Non esiste un team di marketing che, avendo già annunciato delle nuove funzionalità, fa pressione sugli sviluppatori perché le date preannunciate siano rispettate. L'obiettivo è mettere a disposizione del software che funzioni bene, non del software che sia disponibile ad una data prefissata. Anzi, lo sviluppatore ha interesse a rilasciare del software che funzioni bene perché non c'è nessun guadagno a distribuire delle correzioni, a differenza di quanto invece avviene con gli "aggiornamenti" di software chiuso e proprietario.

### 1.3.2. Il Software Libero è più sicuro

Nel software libero non possono esistere backdoor, cioè punti di ingresso conosciuti solo da chi ha sviluppato il codice. Essendo i sorgenti a disposizione di tutti, la presenza di una backdoor verrebbe segnalata in tempi brevi. Ovviamente, tutte le volte che si installa del software libero già compilato non ci può essere alcuna garanzia che uno sviluppatore malintenzionato non abbia inserito del codice non conforme ai sorgenti pubblici. Alla lunga però difetti di questo genere vengono scoperti e possono essere corretti, dato che il codice sorgente originale è sem-

pre a disposizione per essere ricompilato.
È accaduto, per esempio, che le installazioni di un noto Data Base Relazionale proprietario contenessero un problema riguardante la sicurezza degli accessi al Data Base, perché banalmente gli sviluppatori avevano lasciato nel codice una backdoor di cui poi si erano dimenticati. Quando il Data Base in questione è diventato libero ed è stato messo a disposizione di tutti in formato sorgente, l'errore è stato scoperto e corretto nel giro di pochi mesi.

#### 1.3.3. Il Software Libero garantisce la scelta del fornitore

Al di là delle questioni tecniche finora elencate, un altro vantaggio tipico del software libero sta nella libertà di scelta: il cliente può davvero liberamente scegliere il fornitore di software o di servizi migliore disponibile sul mercato.
Forse paradossalmente, la massima libertà di scelta favorisce anche i fornitori stessi. Di primo acchito sembra che con il software libero per un cliente sia più semplice abbandonare un fornitore per un altro, ma questo significa anche che il cliente non è spaventato a dover lavorare con una piccola ditta che egli teme possa sparire in cinque o dieci anni. Molti piccoli sviluppatori software temono di perdere i loro già piccoli guadagni se sviluppassero software libero. In realtà molti programmatori e case di sviluppo vivono dignitosamente scrivendo e vendendo software libero, in un mercato in cui le loro scelte sono realmente libere.
Lo stesso non si può dire di chi basa i suoi programmi su piattaforme software proprietarie: devono costantemente subire "scelte di mercato" prese da altri, spendendo moltissima parte di quanto guadagnano solo per aggiornamenti di licenze per software che non fa quanto promesso dalle brochure pubblicitarie.

#### 1.3.4. Il Software Libero abbassa la barriera all'ingresso

Il software libero per sua natura abbassa le barriere di ingresso al mercato e questo può certamente essere salutato come un lieto evento da parte degli sviluppatori e degli utenti. Gli utenti avranno in tale modo la possibilità di ridurre i costi fissi legati alla loro soluzione salvaguardando quindi una maggiore liquidità all'acquisto di servizi per lo sviluppo di nuove soluzioni. Il fiorire di società che stanno migrando o sviluppando applicazioni su Linux così come l'affermarsi di palmari basati sul kernel Linux lo dimostra. Le imprese libere sono compensate con vendite e profitti per aver legalmente e correttamente soddisfatto gli acquirenti. Il software libero in quanto tale in realtà viene venduto a prezzi talmente bassi che i guadagni più significativi vengono fatti vendendo servizi e assistenza o soluzioni. RedHat, Suse, Caldera, TurboLinux e Mandrake sono alcuni esempi di aziende che si guadagnano da vivere con il software libero.
L'alternativa al software libero è la sorveglianza, ovvero assicurarsi che il software non sia usato o copiato illegalmente. Parlando in generale, la parola "sorveglianza" non è usata - si sente invece parlare di "controllo licenze" o frasi simili.

### 1.4. L'affermarsi di LINUX

Oggigiorno tutte le maggiori società di analisi e consulenza del mercato della Information Technology (IT) riconoscono il ruolo di primaria importanza già acquisito da Linux e pure ne prevedono una sempre maggiore affermazione nei prossimi anni. Al riguardo, una delle maggiori società di analisi del mercato IT, nel luglio 2001, sottolineava come LINUX già rappresentasse con ben il 26,9% di diffusione quale OS utilizzato sui nuovi ambienti server installati nel corso dell'anno 2000 il secondo OS più diffuso (Tabella 1).
Le cifre relative alle spedizioni sono, ovviamente, da intendersi in migliaia di unità.
È importante poi notare come il tasso di crescita annuale previsto per Linux sia notevolmente superiore a quello di qualsiasi altro OS e pertanto una sua sempre maggiore diffusione e rilevanza nei prossimi anni risulta cosa certa.
Per quanto riguarda poi le aree applicative per le quali Linux risulta essere maggiormente utilizzato queste sono (in ordine decrescente per quanto concerne la odierna diffusione): l'area dei Web Server, gli e-mail server, i Network server, i Firewall, i Web Application Server, l'area dello sviluppo delle

| Piattaforma SW | 1999 Spedizioni | 1999 Share (%) | 2000 Spedizioni | 2000 Share (%) | 1999-2000 Crescita (%) |
|---|---|---|---|---|---|
| Windows NT/2000 | 2.086 | 38,4 | 2.508 | 409 | 20,2 |
| **Linux** | **1.322** | **24,3** | **1.645** | **26,9** | **24,4** |
| Netware 3.x,4.x,5.x | 1.064 | 16,9 | 1.030 | 16,8 | –3,1 |
| Combined Unix | 826 | 15,2 | 826 | 13,5 | 0 |
| Other NOS | 140 | 2,6 | 116 | 1,9 | –17,4 |
| Total | 5.437 | 100 | 6.125 | 100 | 12,6 |

**TABELLA 1**
*Confronto tra sistemi operativi in funzione del numero di Server installati (Fonte: IDC)*

applicazioni, l'area dei DataBase Server, del Workgroup ed infine quella dei Transaction Server.
Riassumendo, LINUX si è da sempre più diffuso in nuove aree applicative in linea con la sua crescente evoluzione e crescita tecnologica; ovvero, non appena Linux è risultato sufficientemente maturo per essere utilizzato per le applicazioni legate al mondo di internet/intranet è stato da subito usato per esse ed ora che si è ancora più consolidato tecnicamente è pronto per svolgere anche funzioni di DB Server e Transaction Server e pertanto inizia ad essere ora molto usato anche per tali aree applicative.

## 1.5. Fattori che favoriscono e fattori che contrastano il rapido affermarsi di LINUX

Analisi di mercato hanno mostrato come i fattori che maggiormente stanno contribuendo all'affermarsi di Linux siano: il basso costo iniziale, la sua notevole affidabilità e robustezza, il basso costo a regime, il fatto che rappresenti una alternativa a mondi proprietari, la sua sempre crescente scalabilità , la sua portabilità fra piattaforme HW diverse, la crescente disponibilità di applicazioni ed il fatto che sia open source.
Le stesse analisi di mercato hanno invece evidenziato come i fattori che contrastano l'affermarsi di Linux siano: la mancanza di società che ne assicurino servizio e supporto, la mancanza di skill presso i clienti, la mancanza di applicazioni, il fatto che alcune applicazioni ritenute importanti da parte dei clienti debbano essere ancora migrate a tale piattaforma e la mancanza di cultura e conoscenza da parte del mercato.

## 1.6. Le iniziative di IBM atte a supportare l'utilizzo di LINUX

L'IBM ha fatto proprie le analisi e soprattutto le debolezze sintetizzate ai punti precedenti ed avendo riconosciuto in Linux un OS già adatto oggi per un utilizzo sempre più significativo da parte delle aziende ha deciso, ad inizio 2001, di investire un miliardo di dollari in tale area.
Con tale importante investimento l'IBM ha focalizzato le seguenti aree:
- tutte le piattaforme Server sono state rese in grado di supportare Linux offrendo così ai clienti delle "value proposition" ed opportunità rivoluzionarie assolutamente inimmaginabili fino a solo un anno fa;
- è stata intrapresa la migrazione di moltissimi dei prodotti SW IBM (quale DB2, Websphere, Domino, Tivoli, ecc.) su Linux per le diverse piattaforme HW;
- è stato costituito un cosiddetto Linux Technology Center per il quale lavorano a tempo pieno oltre 250 programmatori IBM e che ha la missione di lavorare assieme alla "Linux Community" per accelerare la maturazione di Linux attraverso lo sviluppo di utilities, tools, codice, ecc.;
- il supporto degli "Open Source Development Laboratories" assieme ad altri leader del mercato IT;
- la costituzione di 11 Centri a livello mondiale che sono in grado di aiutare sia Independent Software Vendor (ISV) sia clienti che volessero fare il porting su Linux di applicazioni attualmente utilizzate su altri OS;
- la costituzione ed il mantenimento di diversi siti web ai quali gli utenti possono collegarsi per ricevere una informazione puntualmente aggiornata relativamente al mondo Linux, al-

le applicazioni disponibili, ai tool a disposizione per gli sviluppatori, ecc.;
- ed infine, l'IBM ha iniziato a fornire con proprio personale una notevolissima quantità di Servizi per i clienti quali servizi di supporto di base ed avanzato, servizi di formazione, di consulenza, di implementazione e di sviluppo di applicazioni.

Per maggiori informazioni ci si può riferire al sito web http://www.ibm.com/linux [5].

### 1.7. Il ruolo delle Distribuzioni

Sebbene Linux sia liberamente disponibile e scaricabile attraverso Internet, lo scaricare il codice sorgente ed il ricompilare lo stesso al fine di farlo funzionare correttamente sulla piattaforma HW posseduta non è una attività così semplice da essere alla portata di tutti. Per tale motivo un discreto numero di società ha iniziato a distribuire Linux. Tali società non cambiano il sistema operativo, che è protetto dalla GPL ma si fanno bensì pagare perché portano Linux sul mercato sotto forma di un set consistente di CD, Manuali e pure assicurano il supporto all'installazione ed all'utilizzo di Linux. IBM ha sottoscritto partnership mondiali con le maggiori società che distribuiscono Linux (quali Red Hat, Suse, Caldera e TurboLinux) e pure alcune partnership locali (con Mandrake, RedFlag, ecc.).

### 1.8. LINUX: aree di utilizzo

Al paragrafo 1.4 abbiamo gia velocemente analizzato le aree applicative per le quali Linux risulta essere oggi maggiormente utilizzato. Tali aree applicative possono essere sintetizzate in **"quattro Modelli Infrastrutturali ed Applicativi"** per i quali Linux può costituire una soluzione veramente innovativa.
Questi quattro modelli sono:
- **Appliances**: ovvero i sistemi dedicati allo svolgimento ottimale di una specifica funzione quali i sistemi che fungono da firewall, web-server, web application server, ecc.;
- **Distributed Enterprise**: Linux è un sistema robusto, affidabile, liberamente distribuibile ed economico ed è quindi perfetto per le aziende che hanno decine, centinaia o migliaia di sedi dove debbono utilizzare le stesse applicazioni con sicurezza, controllo ed in maniera economicamente efficace;
- **Linux Cluster**: ovvero insiemi di server che vengono collegati fra loro al fine di realizzare delle soluzioni infrastrutturali in grado di offrire alte prestazioni e/o superiori livelli di sicurezza a condizioni economicamente interessantissime e *dei quali parleremo più in dettaglio anche nei paragrafi successivi*;
- **Workload Consolidation**: ovvero il processo di razionalizzazione e riduzione del fenomeno di selvaggia ed incontrollata proliferazione di decine, centinaia e addirittura migliaia di server che sta oggi colpendo e mettendo in crisi la infrastruttura IT di tante aziende e Service Provider.

Grazie al fatto che oggi Linux può girare eccezionalmente bene ed assai efficacemente anche sui più grossi server IBM delle famiglie zSeries, iSeries e pSeries, e grazie all'estrema possibilità di portare e migrare applicazioni dai mondi proprietari verso il mondo Linux, i clienti possono pensare di "consolidare" centinaia di server su di un solo sistema con enormi vantaggi dal punto di vista del "Total Cost of Ownership" della soluzione, maggiore semplicità gestionale e riduzione delle risorse richieste.

## 2. CLUSTER LINUX: DESCRIZIONE GENERALE

Kai Hwang e Zhiwei Xu nel loro libro *Scalable Parallel Computing* [4], definiscono i *cluster* come insiemi interconnessi di interi computer (*nodi*) che lavorano congiuntamente come un singolo sistema (*single system image*) per fornire un servizio continuativo (*availability*) e servizi efficienti (performance). I cluster il cui scopo prioritario è quello di garantire la disponibilità del servizio sono noti come *HA cluster*, mentre quelli il cui scopo prioritario è di garantire alte prestazioni sono noti come *HPC cluster*. Nel seguito dedicheremo un paragrafo agli HA cluster e uno agli HPC cluster. In entrambi i casi ci limiteremo a considerare i sistemi basati sul sistema operativo Linux.

### 2.1. CLUSTER Linux per la realizzazione di sistemi e soluzioni Altamente Affidabili

Nel disegno di una soluzione ciò che l'architetto deve considerare prioritario, al fine di soddisfare le esigenze dell'azienda che

adotta quella soluzione, è la *realizzazione di un sistema che eroghi il servizio con prestazioni adeguate, nel momento in cui questo servizio effettivamente è richiesto, ad un costo il più piccolo possibile*. *Prestazioni adeguate* significa che, salvo alcune limitate eccezioni, le aziende non sono interessate ad avere il massimo delle prestazioni possibili, ma solo ad avere prestazioni adeguate alle loro esigenze. Per esempio, se il sistema ERP aziendale ha tempi di risposta per l'attività di inserimento ordini dell'ordine di 1 o 2 secondi, ben difficilmente i sistemi informativi saranno disposti ad investire ingenti somme di denaro per dimezzare il tempo di risposta. *Servizio effettivamente richiesto* significa che, anche se i sistemi informativi gradirebbero avere un sistema che non si guasti mai e non richieda alcuna interruzione di servizio per operazioni di manutenzione, questo non è certo l'obiettivo primario.

In effetti gli obiettivi sono, in ordine decrescente di importanza:

1. *minimizzare il costo (danno) dovuto a interruzioni dell'erogazione del servizio*: ciò significa che le interruzioni pianificate di servizio sono considerate accettabili, benché ovviamente non siano auspicabili;
2. *minimizzare il numero e la lunghezza delle interruzioni del servizio*: anche se le interruzioni di servizio pianificate non hanno un impatto grande come quelle non pianificate, nondimeno creano problemi di gestione, hanno comunque un impatto negativo sull'utenza se il sistema lavora 24 ore al giorno, possono ridurre le performance (buffering subottimale) e richiedono il riavvio dei processi di elaborazione batch;
3. *minimizzare il numero dei guasti*: anche se con tecniche di mascheramento è possibile far si che il guasto non produca un'interruzione di servizio, le componenti guasti devono però essere sostituite, con un conseguente incremento di lavoro per i sistemi informativi.

Notiamo che mentre l'obiettivo 1 è di importanza cruciale sia per gli utenti che per i sistemi informativi, l'obiettivo 2 è maggiormente importante per i sistemi informativi e l'obiettivo 3 è esclusivamente rilevante per i sistemi informativi.

Nel seguito del paragrafo faremo vedere come, facendo leva sulle caratteristiche avanzate della piattaforma IBM eServer xSeries e sulla robustezza del sistema operativo Linux, con un accurato disegno della soluzione sia possibile soddisfare la grande maggioranza delle esigenze aziendali per mezzo del sistema operativo Linux installato su server eServer xSeries. Vedremo in particolare come l'uso di tecnologie di clustering permetta di raggiungere livelli di disponibilità del servizio potenzialmente superiori al 99.9%.

### 2.1.1. Evitare i guasti

Per evitare o minimizzare i guasti è necessario che il disegno di tutte le componenti della soluzione, separatamente prese, sia finalizzato all'obiettivo di costruire componenti affidabili, ma anche e soprattutto che nella fase di disegno si tenga esplicitamente conto delle complesse interazioni fra componenti durante l'esecuzione dei processi elaborativi. Quanto detto concerne ovviamente sia le componenti HW che le componenti SW. L'affidabilità del kernel Linux è stata soggetta a diversi studi che ne hanno dimostrato una stabilità paragonabile a quella dei più blasonati sistemi Unix proprietari. L'affidabilità complessiva del sistema operativo Linux su piattaforma Intel, ossia della combinazione del kernel Linux con il SW fornito a corredo dalle distribuzioni e della piattaforma Intel sottostante, è stato oggetto di un recente studio di D.H. Brown. La classificazione di D. H. Brown delle distribuzioni Linux in comparazione ai principali sistemi Unix relativamente al criterio RAS (acronimo che sta per le tre parole Reliability, Availability e Serviceability), ha visto prevalere SuSE Linux 7.2 sulle altre distribuzioni e ha evidenziato alcuni punti deboli di Linux, o più precisamente di Linux su piattaforma Intel. Va detto però che alcuni di questi limiti sono in fase avanzata di soluzione, quale per esempio il multi-path I/O, mentre altri sono in via di soluzione con l'arrivo dei primi server basati sull'IBM Enterprise X-Architecture (EXA).

Il problema estremamente complesso della qualità del SW e dei processi di debugging, verifica e test attraverso i quali si riduce la percentuale di difetti presenti nel SW non sa-

ranno esaminati in questo articolo. Un'analisi approfondita e aggiornata di questo argomento è contenuta nel numero monografico dell'IBM System Journal su *Software Testing and Verification* [3].

### 2.1.2. Evitare le interruzioni di servizio

Evitare o minimizzare le interruzioni di servizio in presenza di guasti è possibile tramite tecniche di *mascheramento*. Un tipico esempio è l'uso di un *fault-tolerant team* di schede di rete. Se una scheda si rompe, o se si rompe lo switch al quale la scheda è connessa, viene attivata, in modo trasparente per gli applicativi, la scheda di backup. Le tecniche di mascheramento sono ampiamente utilizzare nel mondo Intel per dischi (RAID), schede di rete (team fault tolerant), ventole di raffreddamento, alimentatori. Fino a qualche mese fa rimaneva problematico la protezione attraverso ridondanza di CPU, memoria e sottosistema di I/O. La ridondanza nel sottosistema di I/O (multipath I/O) è ora possibile con il nuovo driver delle schede Qlogic. La protezione della memoria sarà possibile a breve con l'arrivo dei primi eServer xSeries che supportano il memory mirroring con sostituibilità a caldo dei banchi di memoria, una delle pietre miliari dell'Enterprise X-Architecture. Rimane come ultimo e più difficile ostacolo la ridondanza con sostituibilità a caldo delle CPU. Attualmente tutto ciò che si riesce a fare in caso di blocco di una CPU è il riavvio del server con successivo isolamento della CPU guasta e ripresa dell'attività sfruttando le altre CPU (in un sistema multiprocessore).

### 2.1.3. Evitare le interruzioni di servizio quando il sistema è utilizzato

La trasformazione del tempo di arresto (downtime) da non pianificato in pianificato è di grande utilità per la grande maggioranza dei sistemi; fanno eccezione ovviamente i sistemi che devono lavorare 24 ore al giorno per 365 giorni all'anno. Fra le tecniche più promettenti per raggiungere questo scopo vale la pena citare la *Software Rejuvenation*. Numerosi studi sviluppati a partire dalle metà degli anni '90 hanno evidenziato che molti blocchi di sistemi sono dovuti a errori di programmazione che provocano durante l'esecuzione dei programmi fenomeni quali *memory leaks, unterminated threads, memory bloating*, ecc. In linea di principio è ovvio aspettarsi che il miglior modo per affrontare questi problemi sia definire migliori tecniche di controllo del SW e, in caso di problema, di identificare la componente malfunzionante per poterla riparare. Vi sono varie ragioni per le quali questa non è, quanto meno, l'unica via per affrontare il problema. Da un lato se il baco risiede in un SW proprietario per ottenere una *patch* che corregga l'errore può essere necessario aspettare mesi, e questa è in effetti una delle ragioni del successo del modello di sviluppo Open Source. D'altro canto alcuni degli errori possono essere così complessi da rendere l'individuazione della causa estremamente difficile (i cosiddetti Heisenbugs). In tutti questi casi è di grande beneficio per il sistema adottare tecniche di Software Rejuvenation, ossia di riavvio controllato di singoli gruppi di processi o di tutto il sistema operativo. Anche se il riavvio controllato è una prassi che molti amministratori di sistema applicano da molto tempo, la determinazione del momento ottimale per effettuare il riavvio e l'individuazione di metodiche che permettano di prevedere il presentarsi di guasti, sono problemi tutt'altro che banali che sono stati oggetto di numerosi studi sviluppati congiuntamente da IBM e dalla Duke University. Sulla base di questi studi, basati su sofisticate tecniche matematiche (*Stochastic Reward Nets*), è stato sviluppato il modulo del SW IBM Director, offerto gratuitamente su tutti gli eServer xSeries, che implementa il processo di SW Rejuvenation.

### 2.1.4. Come gestire i guasti

Se non si riesce ad evitare il guasto e soprattutto se non si riesce ad evitare che il guasto si presenti nel momento in cui sistema è utilizzato, non resta che disporre di metodiche che rendano minimo il tempo che intercorre fra il guasto e la riparazione del sistema. Vari accorgimenti aiutano a ridurre il tempo necessario per la diagnosi e quello necessario per la riparazione. Ma sia la criticità sempre crescente dei sistemi informativi che la complessità degli stessi, stanno spingendo il mercato all'adozione sempre più ampia di meccanismi che automatizzino il processo di

ripristino del servizio. Rientrano in questa categoria in particolare le metodologie di replicazione e quelle di clustering. Le tecniche di replicazione, utilizzate in file system distribuiti come il GPFS di IBM e in tutti i maggiori DBMS ,giocano un ruolo molto importante, ma per contenere l'esposizione non saranno analizzate in questa sede. Affronteremo invece i cluster per alta affidabilità.
I cluster per alta affidabilità possono essere classificati secondo molti punti di vista. Per presentarne le caratteristiche elencheremo nel seguito alcuni di questi punti di vista e faremo vedere come alcuni di questi cluster si comportano relativamente a questi. La trattazione ovviamente non ha alcuna pretesa di esaustività. Il lettore potrà trovare informazioni più complete nel sito http://linux-ha.org [1].

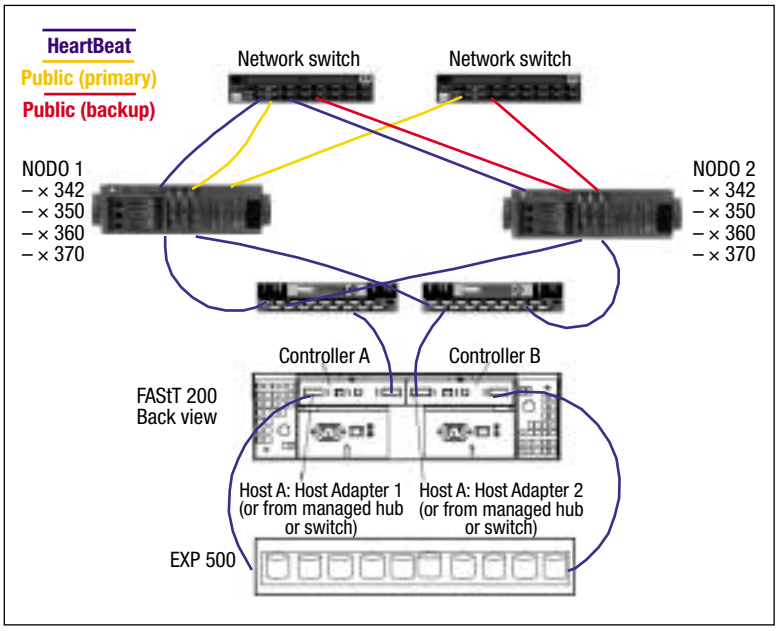



**FIGURA 1**
*Cluster HA linux con IBM eServer xSeries e storage eServer FAStT200*

### L'accesso ai dati

Vi sono sostanzialmente tre categorie di cluster:

- cluster basati sulla totale replicazione dei dati (o clonazione dei server) quali per esempio TurboCluster della TurboLinux o LifeKeeper Data Replication della SteelEye Technologies;
- cluster basati sulla condivisione dei dischi a livello HW ma non a livello SW (*shared nothing*) quali per esempio SteelEye LifeKeeper e Mission Critical DataGuard;
- cluster basati sulla condivisione dei dischi sia a livello HW che ha livello SW (*shared everything*) come per esempio Oracle 9i Real Application Cluster.

La prima categoria di cluster è prevalentemente utilizzata per le web farm o per altre attività a basso carico transazionale. La seconda categoria di cluster è quella con il più ampio spettro di applicazioni. Infatti, può essere utilizzata per proteggere virtualmente ogni applicativo: web server, database server, application server, ecc. La terza categoria di cluster richiede lo sviluppo di un complesso meccanismo per evitare danni all'integrità dei dati dovuti all'accesso simultaneo di più nodi, ossia di un distributed lock manager; per questa ragione non ha mai raggiunto un elevato livello di diffusione. Nel seguito ci concentreremo prevalentemente sulla seconda categoria.

### Il sottosistema dischi condiviso

I nodi possono accedere al sottosistema dischi condiviso per mezzo del protocollo SCSI, del protocollo FC (SAN) o attraverso il protocollo IP (NAS). Di questi il protocollo FC è quello che garantisce la migliore combinazione di affidabilità e performance. Per questa ragione la grande maggioranza dei cluster per alta affidabilità, quantomeno per i server basati su piattaforma Intel, utilizzano la tecnologia FC. I NAS sono un'alternativa interessante per sistemi che non abbiano le esigenze di performance e integrità dei dati che sono tipiche dei sistemi transazionali. La figura 1 rappresenta un cluster Linux per alta affidabilità basato su IBM eServer xSeries e storage IBM FAStT200.

### Le distribuzioni

Un aspetto fondamentale è l'integrazione con le distribuzioni. La tabella 2 fornisce un elenco sintetico delle distribuzioni supportate.

### Gli applicativi

In linea di principio un cluster per alta affidabilità shared nothing può proteggere ogni applicativo "che si comporti bene" ossia che sia in grado di ripristinare la propria attività in seguito ad un'interruzione irregolare. Per esempio un database transazionale può essere protetto con questa tecnica perché il DBMS riporta il sistema in condizioni di inte-

**TABELLA 2**
*Distribuzione supportata dai Software*

| SW | Distribuzioni supportate |
|---|---|
| SteelEye LifeKeeper v. 4.0 | Red Hat 6.x and 7.x, SuSE 7.x, Caldera eServer 2.3, Miracle Linux 1.1 and 2.x e TurboLinux (vedi http://www.steeleye.com/products/linux/system_conf.html per un elenco aggiornato) |
| Mission Critical Convolo Data Guard | Red Hat Linux 6.2, 7.1, SuSE Linux 7.1 e Debian GNU/Linux 2.2r3 (vedi http://www.mclinux.com/products/convolo-requirements.php per un elenco aggiornato) |

grità sfruttando i log. Accanto al problema di quali applicativi possono essere protetti vi è il problema di come proteggere gli applicativi. In generale i SW di clustering forniscono meccanismi generici per controllare lo stato dei processi, spegnerli in caso di malfunzionamento e riavviarli su un altro nodo. L'implementazione specifica dei meccanismi di controllo, avvio e spegnimento dei processi è però demandata ad opportuni script. SteelEye Technologies vende degli Application Recovery Kit che possono essere utilizzati per agevolare il processo di configurazione di applicativi in cluster e un Software Development Kit per consentire lo sviluppo di opportuni script per applicativi per i quali non è disponibile un Application Recovery Kit. L'elenco degli ARK disponibili presso SteelEye Technologies è contenuto nella seguente pagina web http://www.steeleye.com/products/supported_apps.html [7]. Mission Critical fornisce invece supporto per un limitato numero di applicazioni sotto forma di servizio. L'elenco degli applicativi supportati è disponibile all'URL http://www.missioncriticallinux.com/support/supported-applications.php [8].

### 2.1.5. Un Esempio di Linux Cluster per la High Availability (HA)

IBM e Open4.it hanno realizzato un cluster HA e Load Balanced, basato su architettura Pentium, fiber channel e software OpenSource per Linux con le seguenti caratteristiche:

- 40 cpu pentium 4;
- 35 GB ram;
- 1 Tera di storage.

Il cluster è articolato in: una coppia di bastioni in HA, una coppia di balancer in HA, venti server in LB come farm, una coppia di database server in HA, una coppia di publishing server in HA, una coppia di file server in HA, uno storage in fiber channel con controller ridondato, due switch fiber channel e due switch Catalist (Cisco) di rete.
Tale cluster è attuamente in funzione presso Aonet in Milano, dove assicura la continuità dei servizi di web publishing, content management, ecommerce, messaging, groupware web e wap che vengono offerti in modalità ASP.
In particolare le applicazioni che vengono fatte girare sono: Apache, Interchange, Qmail, Courier, Twig, Zope, NFS e vario software aggiuntivo per sicurezza e intrusion detection
Secondo Maruzzelli (Resp. Architettura per Open4.it) "Le ragioni che hanno portato all'acquisto di tale cluster sono il vantaggio competitivo che si realizza da una installazione con cui riescono ad assicurare performance, sicurezza e continuità di servizio ad un gran numero di clienti, con costi decrescenti e con amministrazione unificata. Il vantaggio specifico dell'OpenSource e di Linux risiede invece nell'ampia comunità di sviluppatori software e amministratori che garantiscono un costante aggiornamento delle applicazioni sia in termini di funzionalità che, aspetto molto importante, di sicurezza".

## 2.2. CLUSTER LINUX per la realizzazione di sistemi di High Performance Computing(HPC)

I termini High Performance Computing e High Throughput Computing denotano un ampio gruppo di sistemi caratterizzati da grandi esigenze elaborative. Un elenco anche solo sommario dei settori che fanno uso di cluster per HPC richiederebbe molte pagine. Citiamo a puro titolo di esempio:

- fisica delle alte energie - per esempio il pro-

getto LHC del CERN per la ricerca dell'evanescente bosone di Higgs e la verifica delle teorie supersimmetriche;
- prospezioni geologiche - particolarmente rilevante per le aziende del settore petrolifero;
- studi geologici, in particolare per il problema della previsione dei terremoti;
- Life Sciences - sia la genomica che la proteomica richiedono enormi potenze elaborative per analizzare Petabytes di dati;
- previsioni meteorologiche - per analizzare modelli sempre più dettagliati dell'atmosfera;
- Data mining - per analizzare gli enormi database generati dai sistemi di CRM al fine di evitare di perdere clienti e garantire un livello di servizio più elevato;
- risk management - ad esempio per minimizzare i rischi operativi nei settori finanziari e lungo la supply chain;
- analisi di portafoglio - particolarmente interessante per le società di intermediazione mobiliare che facciano uso di algoritmi statistici o di intelligenza artificiale al fine di prevedere l'andamento dei titoli e/o contenere il rischio di investimenti.

In moltissimi settori nei quali si fa uso di algoritmi per i quali la latenza della comunicazione fra i processi non è critica (*coarse grain parallelism*), i cluster sono il metodo preferenziale per la gestione di elaborazioni complesse. In particolare i cluster basati su Linux si stanno affermando come piattaforma preferenziale per l'High Throughput Computing. Le principali ragioni alla base di questa evoluzione sono:
- il basso costo dei server basati su piattaforma Intel (per esempio gli IBM eServer xSeries);
- l'accresciuta performance erogata dai server Intel, in particolare per i calcoli su interi e più recentemente con il rilascio dei Pentium IV anche per i calcoli in virgola mobile;
- il basso costo del sistema operativo Linux;
- la disponibilità del codice sorgente del sistema operativo Linux che permette di ottimizzarlo in funzione dell'esigenza; citiamo a titolo d'esempio la creazione di sistemi operativi con processi mobili (MOSIX, Qlusters OS);
- la facilità con la quale gli esperti gestori dei sistemi HPC, che hanno prevalentemente una cultura UNIX, possono apprendere l'uso di un sistema Linux;
- le ottime caratteristiche del kernel Linux, particolarmente a partire dal rilascio dei kernel della famiglia 2.4.x.

Notiamo altresì che i tentativi fatti di utilizzare Windows NT/2000 in luogo di Linux non hanno avuto molto successo. Le principali ragioni sono:
- costo del sistema operativo;
- difficoltà di ottimizzazione data l'indisponibilità del codice sorgente;
- sistema di gestione meno flessibile.

### 2.2.1. Beowulf: una delle architetture dei Linux cluster

Introduciamo ora una delle architetture maggiormente utilizzate per la realizzazione di Linux cluster per HPC, dalla definizione di T. Sterling e Don Becker in "How to Build a Beowulf": "*un Beowulf è un gruppo di personal computers (PCs), interconnessi da una qualsiasi tecnologia di rete, su cui girano uno dei sistemi operativi open source con caratteristiche proprie dello Unix*". A titolo esemplificativo, un cluster Beowulf potrebbe essere composto da semplicissimi PC, uno switch di rete che riconosca il protocollo TCP/IP e Linux come sistema operativo. Chiunque abbia una discreta conoscenza della tecnologia Intel, del sistema operativo Linux, e del protocollo TCP/IP può realizzare un Linux cluster.

I server e l'infrastruttura di rete sono però, per quanto importanti, solo due delle componenti di un cluster Linux per HPC. Occorre infatti in primo lugo un applicativo "parallelo", ossia in grado di sfruttare il parallelismo del sistema HW. Inoltre occorre un sistema di gestione efficiente, poiché fino a quando il numero di nodi è piccolo, la manutenzione cluster non è molto difficoltosa, ma quando il numero dei nodi diviene elevato il sistema diventa rapidamente ingestibile in mancanza di un sistema gestione efficiente. In particolare in virtù del progressivo aumento del numero di guasti al crescere del numero di nodi. A tal proposito esistono vari tool sia commerciali che non commerciali che aiutano i system administrator del cluster, per una descrizione di questi tools e delle componenti che fanno parte di un Linux cluster per HPC rimandiamo alla lettura dell'IBM redbook SG24-6041-00 scaricabile all'indirizzo http://www.redbooks.ibm.com [6].

### 2.2.2. Le componenti di Linux cluster per HPC

In termini generali un cluster Linux per High Performance Computing è costituito da

- nodi di elaborazione (per esempio eServer xSeries 330);
- un architettura di rete ad alte prestazioni per la comunicazione fra i nodi (per esempio basata su schede di rete e switch Myrinet) indicata spesso con l'espressione *system area network*;
- un sistema per l'immagazzinamento dei dati;
- sistema operativo Linux (spesso basato sulla distribuzione RedHat);
- compilatori paralleli (Portland Group Fortran F90, C, C++, GNU compilers, Intel F90 compiler);
- librerie parallele (per esempio MPI o PVM);
- Job Management System;
- un sistema di gestione integrato del cluster.

Il tipici nodi utilizzati in un cluster Linux per HPC sono

- server monoprocessori basati sull'architettura IA32 (per esempio eServer xSeries 330);
- server biprocessori basati sull'architettura IA32 (per esempio eServer xSeries 330);
- server quadriprocessori basati sull'architettura IA64 (per esempio eServer xSeries 380).

La scelta del tipo di server dipende essenzialmente dal tipo di applicativo, da considerazioni di costo e da valutazioni sulla complessità di gestione. Notiamo altresì che mentre Linux supporta teoricamente fino a un massimo di 16 processori, il grado di scalabilità al di sopra di 8 processori è tutt'altro soddisfacente con gli attuali kernel. Viceversa, i sistemi basati sull'architettura IA64 rendono possibile l'indirizzamento di grandi quantità di RAM, ma sono attualmente limitati ad un massimo di 4 processori. Un altro punto che occorre tenere presente è che, in un cluster costituito da molti nodi, una percentuale significativa del costo complessivo è rappresentata dai dispositivi a corredo (rack, cablaggio, monitor, ecc.). Al fine di minimizzare questa componente sono utili tecnologie quali l'IBM C2T technology che può consentire una riduzione dei costi dell'ordine di decine di migliaia di dollari.

La *system area network* può essere basata su schede di rete e switch Fast Ethernet, se gli applicativi possono tollerare un'alta latenza nella comunicazione fra processi, da schede e switch Gigabit Ethernet, se è necessaria una più grande larghezza di banda o infine da dispositivi Myrinet, se è fondamentale minimizzare la latenza nella comunicazione fra processi. È importante sottolineare in proposito che in un cluster costituito da decine di nodi l'adozione di un'infrastruttura di rete ad alte prestazioni può far si che il costo della rete eguagli o addirittura superi quello dei server. Una chiara comprensione delle effettive esigenze elaborative è quindi cruciale per contenere i costi.

Il sistema per l'immagazzinamento dei dati può consistere semplicemente dei dischi contenuti nei server o alternativamente di uno o più file server. In questo secondo caso un ruolo particolarmente cruciale è svolto non solo dal tipo di file server, per esempio un IBM Enterprise Storage Server o piuttosto un IBM Network Attached Storage, ma anche dal tipo di cluster file system con il quale si accede a questi dati (per esempio l'IBM General Parallel File System).

Un Job Management System è costituito da tre parti principali:

- uno *user server* che consente agli utenti di sottomettere i job a uno o più code di elaborazione;
- un *job scheduler* che decide come e quando assegnare le risorse ai processi di elaborazione;
- un *resource manager* che effettua l'allocazione delle risorse e controlla l'esecuzione dei job.

Il sistema per la gestione integrata del cluster deve garantire per quanto possibile

- una gestione integrata del cluster che consenta di soddisfare il requisito della *Single System Image*;
- l'installazione automatizzata, controllata e a blocchi della server farm;
- garantire un'elevata fault tolerance.

Per quanto concerne le problematiche di gestione ci limitiamo a citare il prodotto open source xCAT sviluppato da IBM, nonché il blasonato IBM Cluster Management System, recentemente portato su Linux. Relativamente alla fault tolerance è evidente che mentre è cruciale per un cluster costituito da centinaia di nodi minimizzare il costo per nodo ed è spesso accettabile in ca-

so di rottura di un nodo il dover riavviare i job in esecuzione su quel nodo, non è però al tempo stesso accettabile avere un'alta difettosità. Se per esempio un job richiede settimane per essere eseguito, il doverlo rilanciare comporta un ritardo di settimane. Tecniche di Predictive Failure Analysis quali quelli disponibili sugli eServer xSeries possono essere d'aiuto per raggiungere questo obiettivo.

### 2.2.3. Un'alternativa avanzata

Accanto alle metodologie tradizionali che prevedono lo sviluppo di applicativi paralleli per mezzo di librerie quali PVM e MPI si sta diffondendo in vari settori l'alternativa più avanzata di utilizzare la migrazione di processi per ottimizzare la distribuzione del carico fra i nodi di elaborazione. Particolarmente interessante nel mondo Linux è il MOSIX sviluppato presso l'Università di Gerusalemme dal Prof. Amnon Barak del quale è stata recentemente presentata una versione commerciale (Qlusters OS) dall'azienda israeliana Qlusters Inc. Quest'ultimo prodotto contiene due importanti componenti che mancano nella versione open source (MOSIX): uno scheduler e un queue manager.

### 2.2.4. Un Esempio di Linux Cluster per HPC

Un esempio di Linux cluster per HPC in Italia è quello che IBM ha realizzato per il Cineca di Bologna e che ha le caratteristiche riportate in tabella 3.

Tale cluster è oggi in funzione presso il Cineca di Bologna dove è utilizzato per fare girare codici di calcolo relativi alle seguenti aree : astrofisica, chimica computazionale e fisica dello stato solido.

Le ragioni per le quali il Cineca ha deciso di installare un tale HPC Cluster sono dovute alla convinzione che sia gia oggi utile sperimentare le tecnologie di clustering Linux e Beowulf per il supercalcolo a conferma della crescente maturità raggiunta da Linux.

| | |
|---|---|
| Compute node Model: | IBM x330 servers |
| Processor (PE): | Intel Pentium III 1.133 GHz |
| Number of PEs: | 128 |
| Processor per node: | 64 nodes with 2 proc. |
| L2 Cache: | 512 kbit per processor full speed |
| DRAM: | 64 Gbytes (1GB/node) |
| Disk space: | 2.7 TB |
| Peak performance: | 145 Gflop/s |
| OS: | Linux |
| Internal Network: | Miricom LAN Card "C" Version |
| Available compilers: | Portland Group Fortran F90, C, C++<br>GNU compilers<br>Intel F90 compiler |
| Parallel libraries: | MPI |

**TABELLA 3**
*Caratteristiche di un cluster HPC*

## Ringraziamenti

Giunti al termine di tale articolo ci fa piacere il ricordare e ringraziare il Dr. Stefano Maffulli (Associazione Culturale MILUG di Milano) per il prezioso supporto fornitoci e per la passione con la quale promuove il software Open Source, il Dr. Erbacci (CINECA di Bologna) per il supporto offerto per quanto riguarda le informazioni relative alla installazione CINECA ed il Dr. Maruzzelli per le informazioni e considerazioni forniteci relativamente alla installazione Open4.it.

Vogliamo poi concludere con una esortazione ed un incitamento al guardare già oggi a LINUX come una possibilità reale che può offrire un vantaggio competitivo all'interno della vostra azienda, del vostro istituto, della vostra scuola perchè sicuramente vi sono gia oggi moltissime aree della vostra infrastruttura informatica che potrebbero trarre dei vantaggi, a volte impensabili, dall'utilizzo di Linux e di altro codice Open Source.

## Bibliografia

[1] High-Availability Linux Project: HTTP://WWW.LINUX-HA.ORG

[2] History of the OSI: http://www.opensource.org/docs/history.html

[3] IBM: Software Testing and Verification. *IBM System Journal*, Vol. 41, n. 1, 2002. http://www.research.ibm.com/journal/sj41-1.html

[4] Kai Hwang, Zhiwei Xu: *Scalable Parallel Computing: Technology,Architecture,Programming*. McGraw-Hill Companies, 1998. http://shop.barnesandnoble.com/booksearch/results.asp?WRD=hwang+xu&userid=66GToH2D8T

[5] Linux at IBM: http://www-1.ibm.com/linux

[6] RedBooks: http://www.redbooks.ibm.com

[7] Supported Apps: http://www.steeleye.com/products/supported_apps.html

[8] Supported Layered Products: http://www.missioncriticallinux.com/support/supported-applications.php

[9] The GNU Project: http://www.gnu.org/gnu/the-gnu-project.html

GIAMPAOLO AMADORI ingegnere nucleare con specializzazione successiva in marketing, è responsabile IBM delle attività di Vendita e Marketing Linux per la Regione Sud Europea. In IBM, dal 1989, ha ricoperto diversi ruoli di sempre maggiore responsabilità quali: Responsabile attività di Marketing per il Mercato Italiano delle Small and Medium Businesses, Direttore Vendite Italia per i Midrange Server.
E-mail: giampaolo_amadori@it.ibm.com

GIANFRANCO BAZZIGALUPPI è Manager of University Relations di IBM Italia.
Esperto di economia e organizzazione, attivo nella ricerca (IreR; Centro Studi di IBM) e nella docenza universitaria, dal 1994 al febbraio 2001 è stato responsabile della Fondazione IBM Italia.
Giornalista pubblicista, ha diretto dal 1993 al 1999 la rivista If, edita da G. Mondadori.
E-mail: bazzi@it.ibm.com

WALTER BERNOCCHI è un Advisor IT Specialist e lavora in IBM Italia per il gruppo xSeries Pre-Sale Technical Support. Ha iniziato la sua collaborazione con IBM nel 1989. Da marzo 2000 lavora a tempo pieno per sviluppare e supportare soluzioni in ambiente Linux. Ha esperienza di C/C++, Java, SunOS, Solaris e AIX. E' Red Hat Certified Engineer (RHCE).
E-mail: walter_bernocchi@it.ibm.com

MAURO GATTI dottore in Fisica e in Ricerca in Fisica, è team leader dell'IBM EMEA eServer xSeries Solution Architects team. La sua area di specializzazione è il disegno infrastrutturale di e-business server farms con eServer xSeries. In particolare si occupa di progetti di sistemi altamente affidabili, Linux, Grid Computing e ERP. Ha esperienze come free lance per alcune grandi aziende dell'IT (Microsoft, HP, IBM).
E-mail: mauro_gatti@it.ibm.com

# COSCIENZA ARTIFICIALE: MISSIONE IMPOSSIBILE?

**Giorgio Buttazzo**

Potrà mai un computer diventare autocosciente? La coscienza è una prerogativa degli esseri umani o potrebbe svilupparsi in un sistema artificiale altamente complesso? Dipende dal materiale di cui sono fatti i neuroni oppure può essere replicata su un hardware differente? Più che fornire delle risposte a tali interrogativi, questo articolo cerca di porre nuove domande, identificare i problemi e discutere le possibili implicazioni legate allo sviluppo di un sistema artificiale autocosciente.

## 1. INTRODUZIONE

Fin dai primi sviluppi dell'informatica, i ricercatori hanno speculato sulla possibilità di costruire macchine intelligenti capaci di competere con l'uomo. Oggi, considerati i successi dell'intelligenza artificiale e la velocità di sviluppo dei computer, tale possibilità sembra essere più concreta che mai, tanto che molti cominciano a credere che in un futuro non tanto lontano le macchine raggiungeranno e supereranno l'intelligenza umana, fino a sviluppare una mente cosciente. Quando si parla di coscienza artificiale, tuttavia, occorre anche considerare gli aspetti filosofici della questione. Un processo di calcolo che evolve in un computer può essere considerato simile al pensiero? Viceversa, il pensiero di una mente umana può essere considerato un processo di calcolo? La coscienza è una prerogativa dei soli esseri umani? Dipende dal materiale di cui sono fatti i neuroni o può essere replicata su un hardware differente?
Oggi nessuno è in grado di fornire una risposta esauriente a questi interrogativi, e allo stato attuale delle conoscenze possiamo solo imbastire una discussione di carattere interdisciplinare a cavallo di settori quali l'informatica, la neurofisiologia, la filosofia e la religione. Anche la fantascienza, essendo un prodotto dell'immaginazione umana, ha un ruolo importante in questa discussione, in quanto, esprimendo i desideri e le paure umane sulla tecnologia, può influenzare il corso del progresso. In effetti, a livello sociale, la fantascienza agisce fornendo una sorta di simulazione di scenari futuri che contribuiscono a preparare la gente ad affrontare transizioni cruciali o a prevedere le conseguenze dell'uso di nuove tecnologie.

## 2. LA VISIONE FANTASCIENTIFICA

Nella letteratura e nella cinematografia fantascientifica, le macchine pensanti costituiscono una presenza costante. Sin dai primi anni cinquanta, i film di fantascienza hanno ritratto i robot come macchine sofisticate costruite per svolgere operazioni complesse al servizio dell'uomo, per lavorare in ambienti

ostili, oppure pilotare astronavi in viaggi galattici [3]. Nello stesso tempo, però, i robot intelligenti sono stati spesso ritratti anche come macchine pericolose, capaci di cospirare contro l'uomo attraverso piani subdoli. L'esempio più significativo di robot con queste caratteristiche è HAL 9000, protagonista principale del film *2001 Odissea nello Spazio* di Stanley Kubrick (1968). Nel film, HAL controlla l'intera nave spaziale, parla dolcemente con gli astronauti, gioca a scacchi, fornisce giudizi estetici su quadri, riconosce le emozioni dei membri dell'equipaggio, ma finisce per uccidere quattro dei cinque astronauti per perseguire un piano elaborato al di fuori degli schemi programmati [14].
In altri film di fantascienza, come T*erminator e Matrix*, la visione del futuro è ancora più catastrofica: i robot diventano autocoscienti e prendono il sopravvento sulla razza umana. Ad esempio, *Terminator 2 – Il giorno del Giudizio*, di James Cameron (1991), comincia col mostrare una guerra terribile tra robot e umani, mentre sullo schermo compare la scritta:

LOS ANGELES, 2029 A.D.

Secondo la trama del film, dopo che la Cyberdyne era diventata la maggiore ditta fornitrice di sistemi militari intelligenti computerizzati, tutti i sistemi di difesa erano stati modificati per diventare completamente autonomi e supervisionati da Skynet, un potente processore neurale costruito per prendere decisioni strategiche di difesa. Tuttavia, Skynet comincia ad apprendere con una velocità esponenziale e diventa autocosciente 25 giorni dopo la sua attivazione. Presi dal panico, gli ingegneri provano a disattivarlo, ma Skynet reagisce violentemente scatenando un attacco militare contro gli umani [4].
Solo in pochissimi film, i robot sono descritti come degli assistenti affidabili, che cooperano con l'uomo piuttosto che cospirare contro. Nel film *Ultimatum alla Terra*, di Robert Wise (1951), Gort è forse il primo robot (extraterrestre in questo caso) che affianca il capitano Klaatu nella sua missione di portare il messaggio di pace agli umani al fine di evitare la loro autodistruzione. Anche in *Aliens - Scontro Finale* (secondo episodio della fortunata serie, diretto da James Cameron nel 1986), Bishop è un androide sintetico il cui scopo è di pilotare l'astronave durante la missione e proteggere l'equipaggio. Rispetto al suo predecessore (presente nel primo episodio), Bishop non è affetto da malfunzionamenti e rimane fedele alla sua missione fino alla fine del film.
Tra gli altri robot buoni ricordiamo Andrew, il robot maggiordomo NDR-114 nel film *L'uomo bicentenario* (Chris Columbus, 1999) e David, il robot bambino nel film *A.I. Intelligenza Artificiale* (Steven Spielberg, 2001). Entrambi cominciano ad operare come macchine preprogrammate, ma col tempo finiscono per acquisire una coscienza di esistere.
La duplice connotazione spesso attribuita ai robot della fantascienza rappresenta il desiderio e la paura che l'uomo ha verso la propria tecnologia. Da una parte, infatti, l'uomo proietta nel robot il suo irrefrenabile desiderio d'immortalità, materializzato in un essere artificiale potente e indistruttibile, le cui capacità intellettive, sensoriali e motorie sono amplificate rispetto a quelle di un uomo normale. D'altro canto, però, esiste la paura che una tecnologia troppo avanzata (quasi misteriosa per la maggior parte di persone) possa sfuggire di mano e agire contro lo stesso uomo (vedi la creatura del Dr. Frankenstein, HAL 9000, Terminator, e i robot di *Matrix*). Il cervello positronico dei robot di Isaac Asimov deriva dalla stessa sensazione di disagio: esso era il risultato di una tecnologia così sofisticata che, sebbene il processo per costruirlo fosse completamente automatizzato, nessuno conosceva più i dettagli del suo funzionamento [2].
I recenti progressi dell'informatica hanno influenzato le caratteristiche dei robot della fantascienza moderna. Ad esempio, le teorie sul connessionismo e le reti neuronali artificiali (mirate a replicare i meccanismi di elaborazione tipici del cervello umano) hanno ispirato il robot Terminator, il quale non è solo intelligente, ma è in grado di apprendere in base alla sua esperienza. Terminator rappresenta il prototipo del robot che abbiamo sempre immaginato, in grado di camminare, parlare, percepire il mondo, tanto da essere indistinguibile da un essere umano. Le sue batterie

possono fornirgli energia per 120 anni, e un circuito di alimentazione alternativo consente di tollerare i danni al circuito primario. Ma, la cosa più importante è che Terminator può apprendere! Egli è controllato da un processore neurale, ossia un computer che può modificare il suo comportamento in base alle esperienze vissute.
Dal punto di vista filosofico, l'aspetto più intrigante sollevato dal film, è che tale processore neurale è così complesso che comincia ad apprendere a velocità esponenziale fino a diventare autocosciente! In tal senso, il film solleva una questione importante sulla coscienza:

"*Può mai una macchina diventare autocosciente?*"

Prima di affrontare questa questione, tuttavia, dovremmo chiederci:

"*Come possiamo verificare che un essere intelligente sia autocosciente?*".

## 3. IL TEST DI TURING

Nel 1950, il pioniere dell'informatica Alan Turing si pose un problema simile, ma riguardo all'intelligenza. Allo scopo di stabilire se una macchina possa o no essere considerata intelligente come un essere umano, egli propose un famoso test, noto come il test di Turing: un computer e una persona interagiscono con un esaminatore esterno attraverso due terminali. L'esaminatore pone delle domande su qualsiasi argomento utilizzando una tastiera. Sia il computer che l'uomo inviano delle risposte che l'esaminatore legge sui monitor corrispondenti. Se l'esaminatore non è in grado di determinare a quale terminale è connessa la persona e a quale il computer, si dice che il computer ha superato il test di Turing e può essere considerato intelligente come l'umano che è dall'altra parte.
Nel 1990, il test di Turing ha ricevuto il suo primo riconoscimento formale da parte di Hugh Loebner e del Cambridge Center for Behavioral Studies del Massachusetts, che hanno instaurato il premio Loebner in Intelligenza Artificiale [11]. Loebner offre un premio di 100.000 dollari per il primo computer in grado di fornire risposte indistinguibili da quelle di un umano. La prima competizione è stata tenuta presso il Computer Museum di Boston nel Novembre del 1991. Per qualche anno, la gara è stata limitata ad un singolo argomento ristretto, ma le competizioni più recenti, a partire dal 1998, hanno eliminato la restrizione sugli argomenti del colloquio. Ciascun giudice, dopo la conversazione, da un punteggio da 1 a 10 per valutare l'interlocutore, dove 1 significa umano e 10 computer [9]. Finora, nessun computer ha fornito risposte totalmente indistinguibili da un umano, ma ogni anno i punteggi medi ottenuti dai computer tendono ad essere sempre più vicini a 5. Oggi, il test di Turing può essere superato da un computer solo se si restringe l'interazione su argomenti moto specifici, come gli scacchi.
L'11 Maggio 1997, per la prima volta nella storia, un computer chiamato Deep Blue ha battuto il campione del mondo di scacchi Garry Kasparov, per 3.5 a 2.5. Come tutti i computer, tuttavia, Deep Blue non comprende il gioco degli scacchi come potrebbe fare un umano, ma semplicemente applica delle regole per trovare la mossa che porta ad una posizione migliore, in base ad un criterio di valutazione programmato da scacchisti esperti.
Claude Shannon ha stimato che nel gioco degli scacchi lo spazio di ricerca della soluzione vincente comprende circa 10120 possibili posizioni. Deep Blue era in grado di analizzare circa 200 milioni ($2 \cdot 10^8$) di posizioni al secondo. Per esplorare l'intero spazio di ricerca Deep Blue impiegherebbe circa $5 \cdot 10^{111}$ secondi, corrispondenti a $10^{95}$ miliardi di anni, un tempo di gran lunga superiore all'età del nostro universo (stimata sui 15 miliardi di anni). Ciononostante, la vittoria di Deep Blue è attribuibile alla sua velocità di calcolo combinata con gli algoritmi di ricerca utilizzati, in grado di considerare vantaggi posizionali e materiali (è stato stimato che Kasparov elabori le mosse ad una velocità di tre posizioni al secondo). In altre parole, in questo caso la superiorità del computer è dovuta all'applicazione della forza bruta, piuttosto che ad un'intelligenza artificiale sofisticata.

Nonostante ciò, in molte interviste rilasciate alla stampa, Garry Kasparov ha rivelato che in certe situazioni aveva la sensazione di giocare non contro una macchina, ma contro un umano. Spesso ha apprezzato la bellezza di certe soluzioni ideate dalla macchina, come se essa fosse guidata da un piano intenzionale, piuttosto che da una forza bruta. Dunque, se accettiamo la visione di Turing, dobbiamo ammettere che Deep Blue ha superato il test e gioca a scacchi in modo intelligente.
Oltre agli scacchi, vi sono altri settori in cui i computer stanno raggiungendo le capacità umane, e il loro numero aumenta ogni anno. Nella musica, per esempio, esistono molti programmi commerciali in grado di generare linee melodiche, o addirittura interi brani, secondo stili desiderati, che vanno da Bach alla musica jazz. Esistono anche programmi in grado di generare improvvisazioni eccellenti su basi armoniche predefinite, secondo stili di grandi jazzisti come Charlie Parker e Miles Davis, molto meglio di quanto possa fare un musicista umano di medio livello. Nel 1997, Steve Larson, un professore di musica dell'università dell'Oregon, propose una variante musicale del Test di Turing, chiedendo ad una commissione di ascoltare un insieme di brani di musica classica e di indicare quali fossero stati scritti da un computer e quali da un compositore autentico. Il risultato fu che molti brani generati dal computer furono classificati come composizioni autentiche e viceversa, indicando che anche in questo campo il computer ha superato il Test di Turing [10].
Altri settori in cui i computer stanno diventando abili come gli umani includono il riconoscimento del parlato, la diagnosi degli elettrocardiogrammi, la dimostrazione di teoremi, ed il pilotaggio di velivoli. Nel prossimo futuro, tali settori si espanderanno velocemente e comprenderanno attività sempre più complesse, come la guida di automobili, la traduzione simultanea di lingue, la chirurgia medica, la tosatura dell'erba nei giardini, la pulizia della casa, la sorveglianza di aree protette, fino ad includere forme artistiche, finora prerogativa dell'uomo, come la pittura, la scultura e la danza.
Tuttavia, anche ammesso che le macchine diventino abili come l'uomo, o più dell'uomo, in molte discipline, tanto da essere indistinguibili da esso nel senso indicato da Turing, potremmo concludere che esse hanno raggiunto l'autocoscienza? Naturalmente no. Qui la questione diventa delicata, in quanto, se l'intelligenza è l'espressione di un comportamento esteriore che può essere osservato e misurato mediante test specifici, la coscienza di un individuo è una proprietà interna del cervello, un'esperienza soggettiva, che non può essere misurata dall'esterno. Al fine di affrontare questo problema è necessario svolgere alcune considerazioni filosofiche.

## 4. LA VISIONE FILOSOFICA

Da un punto di vista puramente filosofico, non è possibile verificare la presenza di una coscienza in un altro cervello (sia esso umano che artificiale), in quanto questa è una proprietà che può essere osservata solo dal suo possessore. Poiché non è possibile entrare nella mente di un altro essere, allora non potremo mai essere sicuri della sua capacità di essere cosciente. Tale problema è affrontato in modo approfondito da Douglas Hofstadter e Daniel Dennett in un libro intitolato *L'Io della Mente* [8].
Da un punto di vista più pragmatico, comunque, potremmo seguire l'approccio di Turing e dire che una creatura può essere considerata autocosciente se è capace di convincerci, superando alcune prove specifiche. Inoltre, tra gli uomini, la credenza che un'altra persona sia autocosciente è anche fondata su considerazioni di similitudine: poiché tutti noi siamo fatti allo stesso modo è ragionevole credere che la persona che ci sta davanti sia anch'essa autocosciente. Chi metterebbe in dubbio la coscienza del proprio migliore amico? Se invece la creatura di fronte a noi, pur avendo sembianze e comportamenti umani, fosse fatta di tessuti sintetici, organi meccatronici e cervello a microprocessore neurale, forse la nostra conclusione sarebbe diversa.
L'obiezione più comune che spesso viene mossa contro la coscienza artificiale è che i computer, poiché sono azionati da circuiti elettronici che funzionano in modo automatico e deterministico, non possono essere

creativi, provare emozioni, amore, o avere libero arbitrio. Un computer è uno schiavo azionato dai suoi componenti, proprio come una lavatrice, anche se più sofisticato. Il punto debole di questo ragionamento, tuttavia, è che esso può essere applicato anche alla controparte biologica. Infatti, al livello neurale, il cervello umano è anch'esso attivato da reazioni elettrochimiche e ogni neurone risponde automaticamente ai suoi segnali di ingresso in base a precise leggi fisiche. Nonostante ciò, questo modo di operare del cervello non ci preclude la possibilità di provare felicità, amore, ed avere comportamenti irrazionali.

Con lo sviluppo delle reti neurali artificiali, il problema della coscienza artificiale è diventato ancora più intrigante, in quanto le reti neurali tendono a replicare il funzionamento di base del cervello, fornendo un supporto appropriato per realizzare dei meccanismi di elaborazione simili a quelli che operano nel cervello. Nel libro *Impossible Minds*, Igor Aleksander [1] affronta questo argomento con profondità e rigore scientifico.

Se si rimuove la diversità strutturale tra cervello biologico e artificiale, la questione sulla coscienza artificiale può solo diventare di tipo religioso. In altre parole, se si crede che la coscienza umana sia determinata da un intervento divino (diretto o indiretto), allora chiaramente nessuna macchina potrà mai diventare autocosciente. Se invece si crede che la coscienza umana sia una naturale proprietà elettrica sviluppata dai cervelli complessi, allora la possibilità di realizzare un essere artificiale cosciente rimane aperta.

## 5. LA VISIONE RELIGIOSA

Da un punto di vista religioso, l'argomento principale che viene portato contro la possibilità di replicare l'autocoscienza in un essere artificiale è che la coscienza non è una conseguenza dell'attività elettrochimica del cervello, ma un'entità immateriale separata, spesso identificata con l'anima. Questa visione dualistica del problema mente-corpo fu sviluppata principalmente da Cartesio (1596-1650) ed è ancora condivisa da molte persone. Tuttavia, essa ha perso di credibilità nella comunità scientifica e filosofica, poiché presenta diversi problemi che non possono essere spiegati con tale teoria.

- Innanzitutto, se una mente è separata dal cervello di cui fa parte, come può fisicamente interagire con il corpo e attivare un circuito neurale? Quando io penso di muovermi, nel mio cervello avvengono specifiche reazioni elettrochimiche che fanno sì che alcuni neuroni si attivino per attuare i muscoli desiderati. Ma se la mente opera al di fuori del cervello, in che modo essa è in grado di muovere gli atomi per creare impulsi elettrici? Esistono delle forze misteriose (non contemplate dalla fisica) che attivano le cellule neurali? Dobbiamo ammettere che la mente interagisce con il cervello violando le leggi fondamentali della fisica?
- Secondo, se una mente cosciente può esistere al di fuori del cervello, perché noi abbiamo un cervello?
- Terzo, se le emozioni e i pensieri vengono da fuori, perché un cervello stimolato per mezzo di elettrodi e droghe reagisce generando pensieri?
- Infine, perché in pazienti affetti da malattie cerebrali il comportamento cosciente viene seriamente sconvolto dalla rimozione chirurgica di alcune aree del cervello?

Questi ed altri argomenti hanno portato il *dualismo* a perdere credibilità nella comunità scientifica e filosofica. Per risolvere queste inconsistenze, sono state sviluppate molte altre formulazioni alternative sulla questione mente/corpo. Da un lato estremo, il *riduzionismo* rifiuta di riconoscere l'esistenza della mente come esperienza soggettiva privata, considerando tutte le attività mentali come stati neurali specifici del cervello. All'altro estremo, l'*idealismo* rifiuta l'esistenza del mondo fisico, considerando tutti gli eventi sensoriali come il prodotto di costruzioni mentali. Purtroppo, in questo articolo non c'è spazio per discutere esaurientemente tutte le varie teorie sulla questione, e comunque ciò esula dagli scopi di questo lavoro. Tuttavia, è importante notare che, con il progresso dell'informatica e dell'intelligenza artificiale, gli scienziati e i filosofi hanno formulato una nuova interpretazione, in base alla quale la mente è considerata come una forma di calcolo che emerge ad un livello di astrazione più alto rispetto a quello dell'attività neuronale.

## 6. LA VISIONE OLISTICA

Il maggior difetto dell'approccio riduzionistico nella comprensione della mente è di decomporre ricorsivamente un sistema complesso in sottosistemi più semplici, fino ad arrivare ad analizzare e descrivere le unità elementari. Questo metodo funziona perfettamente per i sistemi lineari, in cui le uscite possono essere viste come la somma di componenti semplici. Tuttavia, un sistema complesso è spesso non lineare, pertanto l'analisi delle sue componenti elementari non è sufficiente a comprendere il suo funzionamento globale. In tali sistemi, ci sono delle caratteristiche olistiche che non possono apparire ad un livello inferiore di dettaglio, ma che si manifestano solo quando si considera la struttura globale e le interazioni fra le diverse componenti.
Paul Davies, nel suo libro *Dio e la nuova fisica* [5], spiega questo concetto osservando che una fotografia digitale di un volto è composta da un grande numero di punti colorati (pixel), ciascuno dei quali non rappresenta il volto: l'immagine prende forma solo quando si osserva la foto da una certa distanza che ci permette di vedere tutti i pixel. In altre parole, il volto non è una proprietà dei singoli pixel, ma solo dell'insieme dei pixel.
In *Göedel, Escher, Bach* [7], Douglas Hofstadter espone lo stesso concetto descrivendo il comportamento di una colonia di formiche. Come si sa, le formiche mostrano una struttura sociale complessa e altamente organizzata basata sulla distribuzione del lavoro e sulla responsabilità collettiva. Sebbene ciascuna formica abbia intelligenza e capacità limitate, l'intera colonia mostra un comportamento estremamente complesso. Infatti, la costruzione di un formicaio richiede un progetto sofisticato, ma chiaramente, nessuna formica ha in mente il quadro completo dell'intero progetto. Ciononostante, al livello di colonia emerge un comportamento complesso e finalizzato. In un certo senso, il formicaio può essere considerato un essere vivente.
Sotto diversi aspetti, un cervello può essere paragonato ad un formicaio, poiché composto da miliardi di neuroni che cooperano per raggiungere un obiettivo comune. L'interazione tra neuroni è molto più stretta che tra le formiche, ma i principi di base sono simili: suddivisione del lavoro e responsabilità collettiva. La coscienza non è una proprietà dei neuroni individuali, i quali operano automaticamente come degli interruttori, rispondendo ai segnali di ingresso in base a precise leggi fisiche; ma è piuttosto una proprietà olistica che emerge dalla cooperazione fra neuroni quando il sistema raggiunge un livello di complessità sufficientemente organizzato.
Sebbene la maggior parte di persone non ha problemi ad accettare l'esistenza di caratteristiche olistiche, qualcuno rifiuta di credere che una coscienza possa emergere da un substrato di silicio, ritenendo (non si sa su quale base) che essa sia una proprietà intrinseca dei materiali biologici, come le cellule neurali. Dunque è ragionevole porci la seguente domanda:

"*Può la coscienza dipendere dal materiale di cui sono fatti i neuroni*?"

Paul and Cox, in *Beyond Humanity* [13] dicono: "Sarebbe sbalorditivo se il più potente strumento di elaborazione delle informazioni potesse essere costruito solo a partire da cellule organiche e chimiche. Gli aerei sono fatti con materiali diversi da quelli di cui sono composti gli uccelli, i pipistrelli, e gli insetti; i pannelli solari sono costruiti con materiali diversi da quelli che costituiscono le foglie. Di solito esiste più di un modo per costruire un dato tipo di macchina. [...] Il materiale organico è solo quello con cui la genetica è stata in grado di lavorare. [...] Altri elementi, o combinazioni di elementi, possono rivelarsi migliori allo scopo di elaborare informazione in modo autocosciente."
Se, dunque, sosteniamo l'ipotesi che la coscienza sia una proprietà olistica del cervello, allora la successiva domanda da porci è:

"*Quando un computer diventerà autocosciente?*"

## 7. UNA PREVISIONE AZZARDATA

Tentare di fornire una risposta, se pur approssimativa, alla precedente domanda è senza dubbio azzardato. Ciononostante è possibile determinare almeno una condizione necessaria, senza la quale una macchina non può sviluppare autocoscienza. L'idea si basa sulla semplice considerazione che, per sviluppare

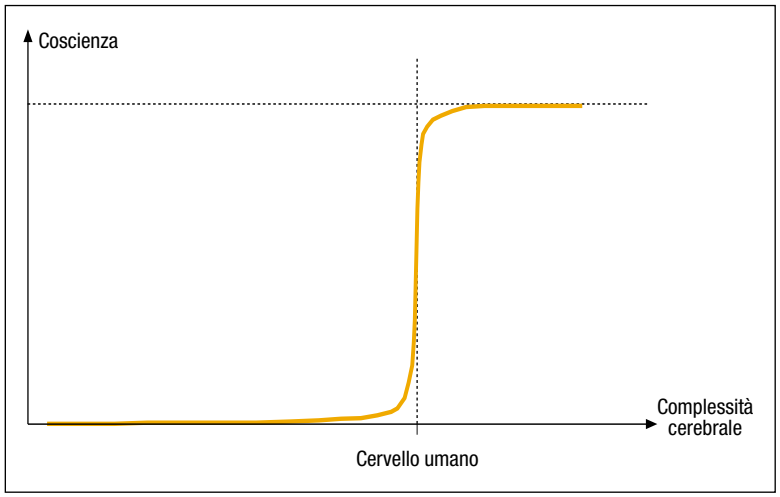



**FIGURA 1**
*Coscienza come funzione della complessità cerebrale*

una forma di autocoscienza, una rete neurale deve essere complessa almeno quanto il cervello umano.
Questa è un'ipotesi ragionevole da cui partire, poiché sembra che i cervelli animali, meno complessi di quello umano, non siano in grado di produrre pensieri consci, nel senso inteso in questo articolo. Dunque, l'autocoscienza sembra essere una funzione della complessità cerebrale a forma di gradino, in cui la soglia è rappresentata dalla complessità del cervello umano (Figura 1).
Ma quanto è complesso il cervello umano? Quanta memoria è richiesta per simulare il suo comportamento con un computer? Il cervello umano è composto da circa mille miliardi ($10^{12}$) di neuroni, e ciascun neurone forma mediamente mille ($10^3$) connessioni (sinapsi) con gli altri neuroni, per un totale di $10^{15}$ sinapsi. In una rete neurale artificiale una sinapsi può essere efficacemente simulata attraverso un numero reale, che richiede 4 byte di memoria per essere rappresentato in un computer. Di conseguenza, per simulare $10^{15}$ sinapsi occorre una memoria di almeno $4\_10^{15}$ byte (4 milioni di Gbyte). Possiamo dunque ritenere che, tenendo conto delle variabili ausiliarie per memorizzare lo stato dei neuroni e altri stati cerebrali, per simulare l'intero cervello umano siano necessari circa 5 milioni di Gbyte. Allora la questione diventa:

"*Quando sarà disponibile tanta memoria in un computer?*"

Durante gli ultimi 20 anni, la capacità della memoria RAM è aumentata in modo esponenziale di un fattore 10 ogni 4 anni. Il grafico di figura 2 illustra ad esempio la tipica configurazione di memoria installata su un personal computer dal 1980 al 2000.
Per interpolazione, si può facilmente ricavare la seguente equazione, che fornisce la dimensione di memoria RAM (in byte) in funzione dell'anno:

$$\text{bytes} = 10^{\left(\frac{\text{year}-1966}{4}\right)}$$

Per esempio, utilizzando l'equazione possiamo dire che nel 1990 un personal computer era tipicamente equipaggiato con 1 Mbyte di RAM. Nel 1998, una tipica configurazione possedeva 100 Mbyte di RAM, e così via. Invertendo la relazione precedente, è possibile predire l'anno in cui un computer sarà dotato di una certa quantità di memoria (assumendo che la RAM continuerà a crescere con la stessa velocità):

$$year = 1966 + 4 \log_{10} (\text{bytes})$$

Duunque, per conoscere l'anno in cui un computer possiederà 5 milioni di Gbyte di RAM, dovremo sostituire tale numero nell'equazione precedente e calcolare il risultato. La risposta è:

**year = 2029**

Un interessante coincidenza con la data predetta nel film Terminator! È interessante anche osservare che una simile previsione è stata derivata indipendentemente da altri studiosi, come Hans Moravec [12], Ray Kurzweil [10], Gregory Paul and Earl Cox [13].
Allo scopo di comprendere a fondo il significato del risultato ottenuto, è importante fare alcune considerazioni. Innanzitutto, è bene ricordare che la data è stata calcolata sulla base di una condizione necessaria, ma non sufficiente, allo sviluppo di una coscienza artificiale. Ciò significa che l'esistenza di un potente computer dotato di milioni di gigabyte di RAM non è sufficiente da solo a garantire che esso diventerà magicamente autocosciente. Ci sono altri fattori importanti che influiscono su questo processo, quali il progresso delle teorie sulle reti neurali artificiali e la compren-

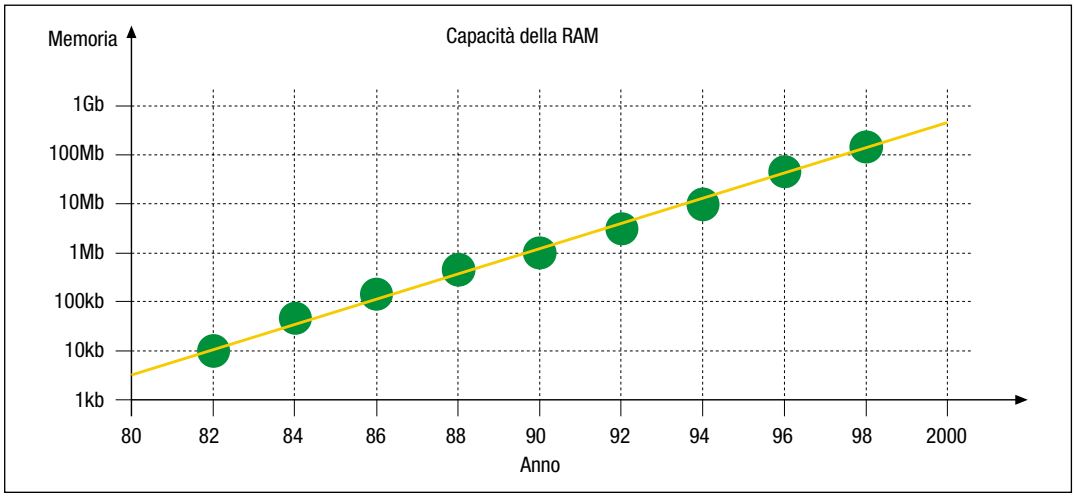



**FIGURA 2**
*Tipica configurazione della RAM (in byte) installata sui personal computer negli ultimi venti anni*

sione dei meccanismi biologici del cervello, sui quali è impossibile tentare una stima precisa. Inoltre, qualcuno potrebbe obiettare che il calcolo effettuato si riferisce alla memoria di un personal computer, che non rappresenta il top della tecnologia nel settore. Qualcun altro potrebbe osservare che la stessa quantità di memoria RAM potrebbe essere disponibile utilizzando una rete di computer oppure dei meccanismi di memoria virtuale che sfruttino lo spazio di uno o più hard disk. In ogni caso, anche utilizzando diversi numeri, il principio di base rimarrebbe lo stesso, e la data verrebbe anticipata solo di qualche anno.

### 7.1. Cosa dire della legge di Moore?

Alcuni obiettano che la previsione del 2029 si basi su una cieca estrapolazione della tendenza attuale di crescita della RAM, senza considerare gli eventi che potrebbero alterare questa tendenza. La tendenza di crescita esponenziale della potenza di calcolo dei computer fu notata nel 1973 da Gordon Moore, uno dei fondatori dell'Intel, il quale predisse che *il numero di transistor nei circuiti integrati sarebbe continuato a raddoppiare ogni 18 mesi fino al raggiungimento dei limiti fisici*. L'accuratezza di tale previsione negli ultimi 25 anni è stata tale da riferire tale osservazione come la "Legge di Moore". Ma fino a quando tale legge continuerà a valere? Le maggiori industrie di circuiti integrati hanno stimato che la Legge di Moore continuerà ad essere valida per altri 15 o 20 anni. Dopo, quando i transistor raggiungeranno la dimensioni di pochi atomi, l'approccio tradizionale non funzionerà più e il paradigma di costruzione dei circuiti sarà destinato a cambiare. Cosa succederà dopo? Forse l'evoluzione dei microprocessori conoscerà la fine intorno all'anno 2020?

Alcuni studiosi, come Ray Kurzweil [10] e Hans Moravec [12], hanno notato che la crescita esponenziale della potenza di calcolo nei computer è cominciata molto prima dell'invenzione dei circuiti integrati, nel 1958, indipendentemente dall'hardware utilizzato. Di conseguenza, la Legge di Moore sui circuiti integrati non descrive il primo paradigma evolutivo, ma il quinto. Ciascun nuovo paradigma ha sostituito il precedente quando necessario. Ciò suggerisce che la crescita esponenziale non si fermerà con la fine della Legge di Moore. L'industria non è a corto di nuove idee per il futuro e si stanno già sperimentando nuove tecnologie, quali il progetto di chip tridimensionali, i computer ottici e i computer quantistici [6]. Dunque, sebbene la Legge di Moore non sarà più valida nel futuro (poiché non si può applicare ai sistemi non basati sul silicio), la crescita esponenziale della potenza di calcolo dei computer probabilmente continuerà per molti anni a venire.

## 8. ALTRI ASPETTI

### 8.1. È preclusa la coscienza alle macchine sequenziali?

Se la coscienza è una proprietà olistica del cervello emergente dal funzionamento collettivo di una struttura neurale altamente

organizzata, e se tale proprietà non dipende dal materiale con cui sono fatti in neuroni, ma dal tipo di elaborazione che essi svolgono, allora è abbastanza ragionevole credere che la coscienza potrebbe anche emergere in una rete neurale artificiale avente una complessità paragonabile o superiore a quella di un cervello umano. Tuttavia, cosa possiamo dire sulla coscienza in computer sequenziali?".

"*Potrebbe mai una macchina sequenziale sviluppare una coscienza?*"

Se la coscienza è il prodotto dell'elaborazione delle informazioni in un sistema altamente organizzato, allora essa non può dipendere dal particolare substrato hardware che realizza il supporto di calcolo. Di fatto, la maggior parte delle reti neurali oggi utilizzate non sono realizzate a livello hardware, ma simulate in un computer sequenziale, che è molto più flessibile (sebbene più lento) di una rete hardware. Qualcuno potrebbe osservare che una simulazione di un processo è diversa dal processo stesso. Chiaramente ciò è vero quando si simula un fenomeno fisico, quale un uragano o un sistema planetario. Tuttavia, per una rete neurale, la simulazione non è diversa dal processo, poiché entrambi sono dei sistemi di elaborazione delle informazioni. Analogamente, la calcolatrice software disponibile sui nostri PC effettua le stesse operazioni della sua controparte elettronica. Dunque, una calcolatrice hardware e la sua simulazione sequenziale sono funzionalmente equivalenti. Pertanto dobbiamo concludere che se una coscienza può emergere da una rete neurale hardware, essa deve necessariamente svilupparsi anche in una sua simulazione software.

## 8.2. La cognizione del tempo

Può la coscienza dipendere dalla velocità degli elementi di calcolo? Non è facile rispondere a questa domanda, ma l'intuizione suggerisce che dovrebbe essere indipendente, in quanto i risultati di un calcolo non dipendono dall'hardware su cui essi sono eseguiti. Ciononostante, la velocità di calcolo è importante per soddisfare i requisiti temporali imposti dal mondo esterno. Se potessimo idealmente rallentare i nostri neuroni uniformemente in tutto il cervello, probabilmente percepiremmo il mondo come in un film accelerato, in cui gli eventi si susseguono ad una velocità maggiore di quella delle nostre capacità reattive. Ciò è ragionevole, in quanto il nostro cervello si è evoluto ed adattato in un mondo in cui gli eventi importanti per la riproduzione e la sopravvivenza sono dell'ordine delle centinaia di millisecondi. Se potessimo accelerare gli eventi dell'ambiente oppure potessimo rallentare i nostri neuroni, non riusciremmo più ad operare efficacemente in tale mondo, e probabilmente non sopravviveremmo a lungo.
Viceversa, cosa proveremmo se avessimo un cervello più veloce? Oggi, una porta logica è un milione di volte più veloce di un neurone: infatti, i neuroni biologici rispondono entro alcuni millisecondi ($10^{-3}$ s), mentre i circuiti elettronici hanno dei tempi di risposta dell'ordine dei nanosecondi ($10^{-9}$ s). Questa osservazione porta ad una domanda interessante: se mai una coscienza emergerà in una macchina artificiale, quale sarà la percezione del tempo in un cervello milioni di volte più veloce di quello umano?
È ragionevole supporre che ad una macchina cosciente il mondo attorno a sé sembri evolversi più lentamente, come un film al rallentatore. Forse la stessa cosa accade agli insetti, che possiedono un cervello più piccolo ma più reattivo e veloce del nostro. Forse, agli occhi di una mosca, una mano umana che tenta di colpirla appare muoversi lentamente, dando ad essa tutto il tempo di volare via comodamente.
Paul and Cox affrontano questa questione nel libro *Beyond Humanity* [13]: " ... un essere cibernetico sarà in grado di apprendere e pensare a velocità elevatissime. Essi osserveranno un proiettile sparato da una distanza moderata, calcoleranno la sua traiettoria e lo eviteranno se necessario. [...] Immaginiamo di essere un robot vicino ad una finestra. Alla nostra mente veloce, un uccello che attraversi il nostro campo visivo sembra impiegare delle ore, e un giorno sembra durare in eterno." È interessante notare che il problema della percezione del tempo nelle menti cibernetiche è stato considerato per la prima volta

nel cinema nel film *Matrix* (Larry & Andy Wachowski, 1999).

## 9. CONCLUSIONI

Dopo tante discussioni sulla coscienza artificiale, viene da porsi un'ultima domanda:

> "*Perché dovremmo costruire una macchina cosciente?*"

A parte problemi etici, che potrebbero influenzare significativamente il progresso in questo campo, la motivazione più forte verrebbe certamente dall'innato desiderio dell'uomo di scoprire nuovi orizzonti ed allargare le frontiere della scienza. Inoltre, lo sviluppo di un cervello artificiale basato sugli stessi principi di funzionamento di quello biologico fornirebbe un modo per trasferire la nostra mente su un supporto più veloce e robusto, aprendo una via verso l'immortalità. Liberati da un corpo fragile e degradabile, e dotati di organi sintetici sostituibili (incluso il cervello), i nuovi esseri rappresenterebbero il successivo passo evolutivo della razza umana. Questa nuova specie, risultato naturale del progresso tecnologico umano, avrebbe la possibilità di esplorare l'universo, sopravvivere alla morte del sistema solare, cercare altre civiltà aliene, controllare l'energia dei buchi neri, e muoversi alla velocità della luce trasmettendo le informazioni necessarie per replicarsi su altri pianeti.
L'esplorazione dello spazio finalizzata alla ricerca di civiltà aliene intelligenti è già cominciata nel 1972, quando la sonda Pioneer 10 fu lanciata per abbandonare il sistema solare con lo scopo preciso di trasmettere nello spazio informazioni sulla razza umana e il pianeta Terra, come un messaggio in una bottiglia nell'oceano.
Tuttavia, come per tutte le più importanti scoperte dell'uomo, dall'energia nucleare alla bomba atomica, dall'ingegneria genetica alla clonazione umana, il problema reale è stato e sarà quello di tenere la tecnologia sotto controllo, assicurando che essa venga usata per il progresso dell'umanità, e non per scopi catastrofici. In questo senso, il messaggio portato dal capitano Klaatu nel film "Ultimatum alla Terra" rimane ancora il più attuale!

## Bibliografia


[1] Igor Aleksander: *Impossible Minds: My Neurons, My Consciousness*. World Scientific Publishers, October 1997.

[2] Isaac Asimov: *I, Robot* (a collection of short stories originally published between 1940 and 1950). Grafton Books, London, 1968.

[3] Giorgio Buttazzo: *Can a Machine Ever Become Self-aware?. In Artificial Humans*, an historical retrospective of the Berlin International Film Festival 2000, Edited by R. Aurich, W. Jacobsen and G. Jatho, Goethe Institute, Los Angeles, May 2000, p. 45-49.

[4] James Cameron, William Wisher: *Terminator 2*: Judgment Day. The movie script, http://www.stud.ifi.uio.no/~haakonhj/Terminator/Scripts/, 1991.

[5] Paul Davis: *God and the New Physics*. Simon and Schuster Trade, 1984.

[6] Linda Geppert: Quantum transistors: toward nanoelectronics. *IEEE Spectrum*, Vol. 37, n. 9, September 2000.

[7] Douglas Hofstadter: *Göedel, Escher, Bach. Basic.* Books, New York, 1979.

[8] Douglas R. Hofstadter, Daniel C. Dennett: *The Mind's I*. Harvester/Basic Books, New York 1981.

[9] Marina Krol: Have We Witnesses a Real-Life Turing Test?. *IEEE Computer*, Vol. 32, n. 3, March 1999, p. 27-30.

[10] Ray Kurzweil: *The Age of Spiritual Machines*. Viking, 1999, p. 160.

[11] Home page of the Loebner Prize, 1999 (Current 28 January 1999): http://www.loebner.net/Prizef/ loebner-prize.html.

[12] Hans Moravec: *Robot: Mere Machine to Transcendent Mind*. Oxford University Press, 1999.

[13] Gregory S. Paul, Earl Cox: *Beyond Humanity: CyberEvolution and Future Minds*. Charles River Media, Inc., 1996.

[14] David G. Stork: *HAL's Legacy: 2001's computer as dream and reality*. Edited by David G. Stork, Foreword by Arthur C. Clarke, MIT Press, 1997.



GIORGIO BUTTAZZO è Professore Associato di Ingegneria Informatica presso l'Università di Pavia, dove svolge attività di ricerca nei settori dei sistemi in tempo reale, della robotica avanzata e delle reti neurali artificiali. Nel 1987, ha conseguito un Master in Computer Science presso l'Università della Pennsylvania e nel 1991 il Dottorato di Ricerca presso la Scuola Superiore S. Anna di Pisa.
E-mail: buttazzo@unipv.it

# WEB LEARNING: ESPERIENZE, MODELLI E TECNOLOGIE

**Alberto Colorni**

Nell'articolo vengono presentate alcune regole per orientarsi nello sviluppo (ma anche nella valutazione) di progetti di e-learning, viene proposto un sistema di classificazione basato su tre elementi, vengono fornite alcune indicazioni sulle tecnologie e sugli strumenti più consueti. Il lavoro nasce dall'esperienza dell'autore nel campo dell'e-learning, sia nella veste di docente al Politecnico di Milano che in quella di direttore del centro per l'innovazione didattica dello stesso Ateneo. L'articolo si conclude con una breve disamina delle principali obiezioni sulla formazione on-line e con la proposta di alcune idee per superarle.

## 1. SCUOLA PER CORRISPONDENZA O ALTRO ?

Il concetto di spartiacque mi ha sempre affascinato: due gocce di pioggia che sono inizialmente distanti meno di un millimetro finiscono l'una nell'Oceano Indiano e l'altra nel Mar Baltico. Può accadere qualcosa del genere anche per l'e-learning: con poche differenze si potrebbe favorire un ulteriore (e definitivo?) degrado dell'istruzione o promuovere un'opportunità nuova e addirittura stimolante.
In questo articolo cercherò di spiegare le ragioni per cui è possibile e sensato fare dell'online learning senza troppe rinunce. Lo farò avendo in mente alcune obiezioni che in questi anni ho sentito spesso fare e che sono esposte molto bene (insieme a molte altre, su cui peraltro concordo) da C. Stoll [18]: obiezioni sulla mancanza di contatto diretto, sui costi sproporzionati, sul pericolo di semplificazione, sull'appiattimento, in una parola su tutto ciò che si perderebbe con l'adozione di questi mezzi.
Vorrei peraltro evitare il trionfalismo di chi pensa a questi strumenti come a una grande rivoluzione che ci renderà tutti pronti per un futuro "digitale" in cui la potenziale libertà di accesso a milioni di dati produrrà istruzione e cultura (e non, come è invece possibile, un faida-te curricolare che ne è l'esatto opposto).
Ciò che voglio (di)mostrare lo riassumo fin d'ora. Si tratta di poche cose, quasi dei consigli per gli acquisti:

- è opportuno farsi guidare dalle esigenze degli utenti e non dalle tecnologie;
- è necessario lavorare (e molto) sull'interattività e sui vari tipi di feedback;
- è fondamentale progettare un percorso, che offra oltre ai materiali anche dei servizi, delle modalità d'interazione, una responsabilità precisa nella conduzione.

Sviluppo le mie considerazioni da un punto di osservazione privilegiato: sono infatti il direttore del Centro METID del Politecnico di Milano [5] che ha prodotto (o partecipato a) molte esperienze nel settore e sono anche il coordinatore del primo progetto di laurea on-line italiano [12].
Per ragioni di spazio userò un approccio abbastanza schematico, rimandando eventuali

ulteriori approfondimenti a un dibattito successivo, magari on-line.
Inizio quindi subito proponendo una distinzione, che è anche indicativa di tre fasi storiche, tra:
**i** distance learning, in cui si è affrontato l'elemento territoriale;
**ii** e-learning, in cui si è inserita la tecnologia informatica;
**iii** on-line learning (o web learning), in cui diventa importante l'interattività.
Naturalmente ogni modalità comprende le precedenti. In questo lavoro mi occupo (quasi esclusivamente) dell'ultima.

## 2. AI VERTICI DI UN CUBO

Normalmente nella Formazione a Distanza (FaD) si segnalano le rotture di continuità rispetto allo spazio e al tempo. Vorrei qui introdurre un terzo elemento su cui ragionare: la modalità di relazione tra i soggetti coinvolti.
Definisco così i tre assi di uno "spazio della formazione", all'interno del quale possono trovar posto alcuni modelli didattici, più o meno noti, di FaDoL (Formazione a Distanza on-line), come mostrato in figura 1. Considerando una coppia di opzioni rispetto a ciascun asse, si ottiene la situazione che segue:

- spazio → presenza (locale) vs. distanza (remoto);
- tempo → reale (sincrono) vs. differito (asincrono);
- relazione → gerarchica (broadcast) vs. collaborativa (interazione, rete).

In ognuno dei tre casi, la seconda opzione introduce una questione rilevante: rispetto allo spazio la questione del canale di trasmissione, cioè delle tecnologie; rispetto al tempo la questione del ritmo di apprendimento e quindi del monitoraggio; rispetto alla relazione la questione dei ruoli ma soprattutto degli attori del processo formativo.
Seguendo la numerazione della figura, illustro sinteticamente le varie situazioni ottenute combinando le opzioni sopra esposte.

Caso 1: è la ***didattica classica***, basata sull'unità di luogo e di tempo e su un solo vero attore (il docente); le versioni più recenti di questo modello offrono supporti come collegamenti Internet e visione di CD-rom ma non ne cambiano la sostanza.

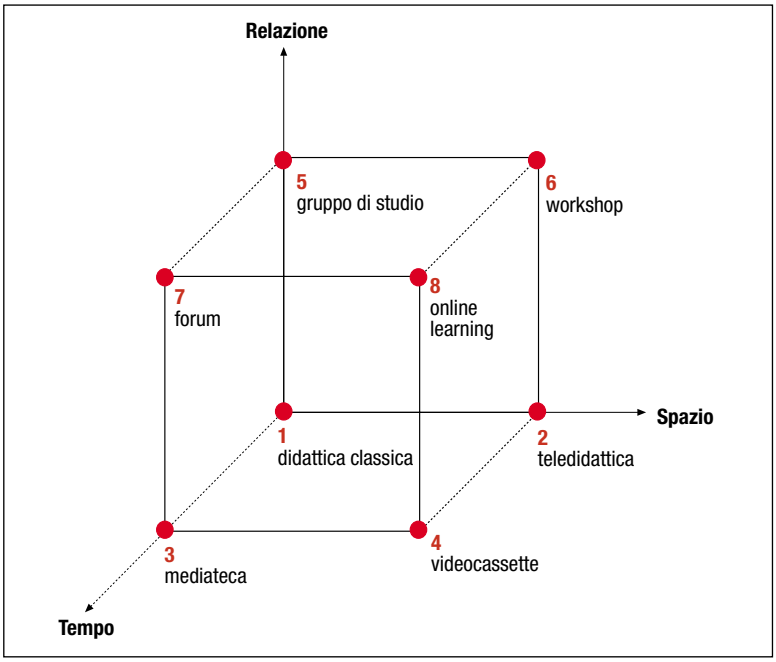



**FIGURA 1**
*Lo "spazio della formazione"*

Caso 2: è la situazione della ***teledidattica*** (e prima ancora della didattica televisiva), con una o più aule remote collegate in tempo reale con l'aula principale dove opera il docente; al Politecnico corsi di questo genere sono regolarmente attivi dal 1993 tra le varie sedi dell'Ateneo.

Caso 3: è la versione multimediale della biblioteca, cioè una ***mediateca*** in cui sono consultabili prodotti multimediali (in genere offline) a supporto dei contenuti indicati dal docente.

Caso 4: è la didattica basata su ***videocassette*** (più recentemente su ***CD-rom***), in cui è possibile vedere o rivedere il docente che fa la sua lezione ripreso da una telecamera (in Italia uno dei casi più noti riguarda il Consorzio Nettuno [8]); è anche il principale modello per la formazione aziendale classica, privilegiando il lavoro del singolo e l'approccio detto ***CBT*** (Computer Based Training).

Caso 5: è la situazione del ***gruppo di studio***, più o meno attrezzato sul piano multimediale e delle connessioni web.

Caso 6: è la tipica situazione del ***workshop*** o della videoconferenza aziendale; i vari partecipanti possono/debbono condividere alcuni strumenti; è quindi necessaria una "piattaforma" che consenta la condivisione.

Caso 7: non c'è un modello rilevante per questa situazione (come sempre, uno schema forza un po' le cose ...); quello che forse si avvicina di più è il ***forum***, con lo scambio asincrono di messaggi.

Caso 8: è il caso che ci interessa maggiormente, l'***on-line learning***; in esso i soggetti coinvolti sono, come vedremo, molti e diversi, con differenti localizzazioni e funzioni; le modalità d'interazione sono assai importanti e di conseguenza è fondamentale la scelta della strumentazione (in particolare della piattaforma), che deve essere fortemente legata alle esigenze che si vogliono supportare.

I casi 1/2/4/8 sono i più interessanti, perchè testimoniano l'evoluzione temporale e gli sforzi fatti per inserire i nuovi strumenti (la TV, il computer, le reti, la multimedialità) nei processi didattici, a partire dagli anni Sessanta fino ad oggi. Su questi argomenti e sulle tecnologie didattiche in generale si può consultare il sito dell'ITD-CNR [11].
Lo scopo dello schema di figura 1 è di mettere in luce il fatto che quando si parla di nuove tecnologie per la didattica si possono intendere progetti e situazioni tra loro molto diversi: non ha quindi senso esprimere posizioni generali, di appoggio o di dissenso, senza riferirle ad un preciso contesto.
In sintesi direi che la principale caratteristica dei modelli (e sono più d'uno) di web learning sta nell'articolazione degli attori, più che nelle possibilità tecnologiche: in particolare la figura del tutor appare cruciale, forse più ancora di quella del docente.

## 3. QUALCHE REGOLA PER ORIENTARSI

Le considerazioni che seguono sono derivate dalla mia esperienza nella gestione del progetto "Laurea on-line del Politecnico" (LLP), sviluppata in collaborazione con Somedia (società del gruppo Espresso-Repubblica) e le cui caratteristiche principali sono descritte in [7], ma anche dalle varie attività del Centro METID a supporto dei docenti dell'Ateneo nei loro progetti di innovazione didattica[1].
Mi riferisco qui a processi didattici di tipo universitario e post-universitario (per esempio master) più che alla formazione aziendale: quest'ultima mi appare prevalentemente individuale e mirata all'ottenimento di "informazioni" per la gestione di situazioni, mentre la prima mi sembra più legata all'acquisizione di "modelli mentali" attraverso percorsi di apprendimento collettivo (la classe). Naturalmente un'integrazione tra i due modelli sarebbe utile a entrambi: offrirebbe alla formazione universitaria una robusta iniezione di pragmatismo e di "customer care", a quella aziendale una solida intelaiatura di obiettivi generali e di tappe parziali.
In ciò che segue è bene tenere a mente una distinzione tra didattica completamente on-line e didattica "mista" (in presenza e on-line). Nel primo caso c'è da aggiungere agli altri il problema della valutazione dello studente e della certificazione dei contributi da lui inviati, mentre nel caso misto ciò appare secondario.
Per le consuete esigenze di sintesi indicherò ora quelli che mi sembrano i nodi principali da sciogliere (nel caso dell'on-line learning), dedicando a ciascuno di essi solo poche righe.
❑ ***Responsabilità***. È importante definire una responsabilità precisa e univoca. Di fronte al sorgere delle ormai moltissime "agenzie" che promettono formazione con tempi brevissimi e modi facili, la questione centrale è quella di far capire che c'è chi governa il processo, assumendosi nei confronti degli studenti la responsabilità dei percorsi formativi, del controllo dell'apprendimento, della valutazione dei risultati. Un processo didattico non può mai essere "facile" (e non deve prometterlo). Esso deve dotarsi di un insieme di organismi che, con differenti compiti e livelli di responsabilità, rispondano del progetto e delle scelte didattiche. Inoltre, per un progetto di questo tipo ci vuole una notevole dose di entusiasmo e di gusto per l'innovazione, da quella didattica a quella tecnologica, da quella gestionale a quella relazionale: serve quindi un "motore". Il risultato dipende (anche) da questo.

[1] In particolare: lezioni in teledidattica tra le 7 sedi lombarde del Politecnico, produzione di CD-rom, messa on-line di materiali, web-conferences, un portale per il supporto alla didattica in presenza, un sistema per la fruizione delle lezioni da parte dei disabili, erogazione di master e corsi brevi, collaborazioni italiane ed europee.

## INGEGNERIA INFORMATICA ON-LINE AL POLITENICO DI MILANO

### Il primo corso di Laurea On-line in Italia

La Laurea in Ingegneria Informatica On-line è un percorso di studio gestito completamente a distanza attraverso il supporto di Internet. Il corso è perfettamente equivalente a quello tradizionale, con le stesse materie di studio, gli stessi crediti formativi, gli stessi carichi di lavoro, gli stessi esami e riconoscimenti della laurea "in presenza" ed è tenuto anch'esso da docenti del Politecnico di Milano. In più, il formato didattico on-line offre agli studenti la possibilità di:

- specializzarsi nell'ICT (Information Communication Technology), utilizzando per lo studio strumenti e mezzi dell'ICT stessa
- personalizzare tempi e modalità di studio senza i vincoli della presenza in aula
- scegliere un percorso di studi su misura a seconda del carico di lavoro che si è in grado di sostenere
- comunicare direttamente via Internet e via mail con tutor e docenti
- socializzare con facilità ed elevata frequenza con gli altri studenti della classe virtuale di appartenenza.

### Piano di studi standard (60 crediti/anno)

Gli studenti devono sostenere circa 8 esami all'anno (4 per semestre), per una totalità di 27 esami in 3 anni, al termine dei quali potranno scegliere se continuare gli studi e frequentare i 2 anni di "specializzazione" oppure se fermarsi alla laurea di primo livello.

| 1° ANNO |
|---|
| **Titolo dell'insegnamento** |
| Elementi di Analisi matematica (A) e di Geometria |
| Fisica sperimentale 1 |
| Fondamenti di Informatica 1 |
| Chimica A |
| Analisi matematica B (per il settore dell'informazione) |
| Fondamenti di Informatica 2 |
| Fisica sperimentale 2 |
| Economia e organizzazione aziendale B |
| **2° ANNO** |
| Elettrotecnica A |
| Fondamenti di automatica (per il settore dell'informazione) |
| Fondamenti di telecomunicazioni |
| Prova di lingua straniera |
| Fondamenti di elettronica |
| Calcolo delle Probabilità |
| Fondamenti di ricerca operativa B |
| Ingegneria del software |
| Attività integrative di laboratorio o progetto |

| 3° ANNO |
|---|
| Automazione industriale |
| Calcolatori elettronici |
| Gestione aziendale B |
| Basi di dati 1 |
| Insegnamento a scelta (area Internet) ([3]) |
| Attività integrative di laboratorio o progetto ([2]) |
| Ingegneria e tecnologia dei sistemi di controllo |
| Insegnamento a scelta (area telecomunicazioni) ([3]) |
| Insegnamento a scelta ([3]) |
| Tirocinio e prova finale |
| ([2]) Le attività integrative di laboratorio o progetto sono associate a un Corso, su proposta dello studente. |
| ([3]) I corsi tra cui lo studente potrà scegliere saranno definiti successivamente. |
| **Totale crediti nei 3 anni: 180** |

### Come funziona

Il Politecnico di Milano, istituzione universitaria tra le più prestigiose in Italia e da sempre all'avanguardia per contenuti e metodi di insegnamento, e Somedia, società privata da anni impegnata nell'organizzazione e gestione di corsi di formazione, tradizionale e on-line, hanno realizzato con la Laurea On-line un formato didattico che integra in modo coerente tutti gli aspetti chiave di un'esperienza di formazione on-line:

- didattica
- tecnologia
- organizzazione.

Allo studente della Laurea On-line viene proposto di partecipare alle attività di una classe virtuale di circa 20-25 studenti seguita da un docente e da un tutor per ciascuna materia. Le principali attività proposte alla classe sono:

**1.** studio individuale (sviluppato secondo il ritmo suggerito dai docenti) basato su:
CD-ROM che propongono i contenuti di base di ciascun insegnamento in versione multimediale
materiali on-line:
SAI (Software di Autointerrogazione)
esercizi commentati
risorse specifiche messe on-line dai docenti durante il corso (slide, spiegazioni, appunti)

**2.** *forme di apprendimento collaborativo on-line*, insieme ad altri studenti, tutor e docenti utilizzando:
strumenti asincroni (forum, e-mail, ecc.)
strumenti sincroni (sessioni live)

**3.** *prove in itinere on-line* periodiche su ciascuna materia.
Alla fine di ogni semestre lo studente dovrà sostenere gli esami in presenza presso la sede di Como del Politecnico di Milano.

❑ ***Progettazione continua***. La formazione on-line coinvolge una molteplicità di soggetti (docenti, tutor, responsabili tecnici, editor multimediali, gestori di rete, ecc.), sul coordinamento dei quali si basa il successo dell'operazione: è necessario raggiungere un certo grado di omogeneità nell'offerta didattica, pur salvaguardando la specificità dei contenuti. La modalità di erogazione (cioè tempi, verifiche, monitoraggio) è un altro elemento chiave. È utile, e spesso necessario, dedicare una struttura specifica a questo scopo, con il compito preciso di governare la complessità del progetto anche attraverso il dialogo con gli studenti. Serve un piano d'azione chiaro, ma anche modificabile in base ai feedback forniti dal sistema (cioè da studenti, docenti, tutor). L'informazione dev'essere costante e ricca (e qui si dovrebbe aprire il problema della reportistica), deve riguardare sia il gruppo che i singoli (permettendo allo studente di avere suoi momenti di autovalutazione), deve essere analizzata verticalmente (su un certo insegnamento) ma anche orizzontalmente (sull'andamento generale dello studente). Per usare una definizione cara agli specialisti dei sistemi di controllo, si tratta di un sistema di gestione "in anello chiuso", in cui l'attività viene corretta secondo i feedback che arrivano. Questa modalità è efficace solo se dispone di continue informazioni sullo stato del sistema: è però la sola che consenta un reale controllo.

❑ ***Apprendimento collaborativo***. È un aspetto cruciale, che distingue il CBT dal web learning. Si tratta di esaltare i momenti di interazione dello studente con i docenti, con i tutor, ma anche con gli altri studenti: ciò che nei videogiochi è la ricerca dell'interattività diventa qui la necessità dell'interazione. Per questo è opportuno creare delle "classi virtuali" di 20-25 studenti, ognuna seguita da un tutor per ogni materia, con momenti di lavoro collettivo sia di tipo asincrono (progetti comuni) che di tipo sincrono (sessioni "live") a cadenza predefinita. Gli studenti di un progetto di web learning sono, in genere, molto motivati e disposti a collaborare: dopo un po' è probabile che ci sia uno sviluppo di queste comunità virtuali rispetto all'obiettivo comune di aiutarsi (non a copiare, ma a risolvere i mille frangenti organizzativi e tecnologici che si presentano ogni giorno). Allo scopo di individuare tempestivamente e seguire gli studenti in difficoltà può essere utile l'attivazione di una figura di "tutor generale".

❑ ***Ritmo*** (***flessibile ma fissato***). Uno dei vantaggi della formazione on-line rispetto a quella in presenza è la quasi totale mancanza di vincoli logistici (orari e aule): ciò permette di proporre delle "finestre" nelle quali ciascun utente può collocare i suoi momenti di fruizione. Ma deve essere, nuovamente, un processo controllato. Per esempio, nel caso della LLP viene messa in linea una "agenda" settimanale in cui i docenti segnalano le parti da studiare e inseriscono gli appuntamenti. La flessibilità consente poi di definire percorsi dedicati: con il progetto LLP il Politecnico ha inaugurato alcuni piani di studio part-time, pensati proprio per quegli studenti (e sono la maggioranza) che non vogliano abbandonare i loro impegni di lavoro.

❑ ***Materiali***. È un discorso cruciale. Non basta scaricare dalla rete materiali già esistenti, è necessario riprogettarli: normalmente, infatti, i materiali didattici disponibili in rete sono utili come supporto del docente in aula, quindi come materiali integrativi. Se invece parliamo di un percorso interamente on-line è necessaria la definizione di formati e tempi di erogazione autonomi. Il docente definisce lo story-board, ma è l'editor multimediale la figura su cui grava il compito di rendere i contenuti fruibili: ruolo non esecutivo ma realmente progettuale, che richiede di operare a stretto contatto con il docente interpretandone le intenzioni e prevedendone gli utilizzi. Inoltre, è utile fornire input differenti: si tratta di sfruttare tutto il potenziale della multimedialità, dalla valorizzazione della rete (prevedendo percorsi guidati dal docente) alla integrazione di vari formati didattici (on e offline). Nel caso della LLP gli input allo studente seguono tre strade diverse, con reciproci "rinforzi": dei CD-rom inviati preventivamente, dei materiali messi in rete periodicamente, delle sessioni "live" settimanali in cui la classe intera discute con il docente o il tutor. I continui rimandi sono funzionali all'idea, ormai largamente accettata, che modalità diverse di esprimere lo stesso contenuto siano più funzionali all'apprendimento.

❑ ***Approccio***. Il "learning by doing" viene spesso citato nei progetti di web learning. Su questo punto ci sono però le difficoltà mag-

giori, per il fatto che i contenuti sono forniti da docenti che (con qualche eccezione) considerano più importante l'insegnamento dell'apprendimento: l'attenzione è quindi sulle cose che il docente ritiene importanti più che sul come esse vengono assimilate dallo studente. Gli sforzi fatti iniziano a produrre i loro frutti, mutuando dalla formazione aziendale modelli didattici basati sull'utente. Il pericolo però è quello di proporre grandi studi di casi e sistemi "fai da te", mentre il problema è, al solito, il controllo del processo.

❑ ***Monitoraggio***. Il controllo sulla qualità di un progetto di formazione on-line pone la questione di come valutare un servizio di questo tipo. La LLP ha scelto come organismo di valutazione un ente esterno[2]. La valutazione può essere fatta da vari punti di vista: l'efficacia didattica, quella tecnica, quella comunicativa (mentre solitamente nella fase iniziale di un progetto, nella quale prevalgono gli investimenti, è più difficile valutare l'efficienza). In questo senso, la collaborazione con un partner esterno come Somedia (che pure ha lasciato totale autonomia al Politecnico nelle scelte didattiche) ha offerto un elemento di stimolo su due fronti: quello dell'analisi delle esigenze dell'utenza e quello dell'attenzione a elementi di efficienza (dichiarando obiettivi, tempi, modalità) di solito non presenti nella formazione universitaria.

Ho detto all'inizio che è opportuno farsi guidare dalle esigenze degli utenti e non dalle tecnologie. Ecco qualche altro esempio di come alcune delle caratteristiche sopra indicate possano essere inserite all'interno di progetti di web learning (mi riferisco naturalmente a progetti che conosco bene, cioè a progetti METID).
Il sistema Corsi on-line [5], offerto ai docenti del Politecnico come supporto alla loro didattica in presenza, è un esempio della necessità di fornire, oltre che materiali in diversi formati, anche, ma dovrei dire soprattutto, servizi ai docenti: per la gestione della classe, per l'aggiornamento dei contenuti, per la comunicazione verso gli studenti, per il testing della loro preparazione. La figura 2 mostra una delle pagine


**FIGURA 2**
*Una pagina del sito Corsi on-line di METID*

del sistema Corsi on-line: nella barra di sinistra è presentata l'organizzazione dei servizi.
Il progetto FormAmbiente [9] ha l'obiettivo di addestrare funzionari regionali all'utilizzo di pacchetti software su tematiche ambientali. È organizzato con materiale in rete e momenti in presenza (corsi brevi della durata di qualche giorno). Qui l'aspetto importante è dato dalla possibilità di definire diversi "profili di utente", ciascuno dei quali ha differenti materiali e servizi di cui può usufruire. Ciò consente sia il controllo degli accessi che la differenziazione dei percorsi tra gli utenti (il discorso dei percorsi differenziati si collega a una delle critiche di Stoll su cui tornerò nelle conclusioni).
In questi esempi, come si attua l'interazione? Nel caso della LLP, essa avviene con le sessioni "live", i test on-line e le prove in itinere, la web-conference (con la possibilità di porre domande in tempo reale), l'uso delle mail, gli scambi di elaborati tra studenti e tutor, i progetti di corso (a volte fatti in gruppo); inoltre nei CD-rom sono previsti momenti di autovalutazione (le cosiddette "piazzole di sosta") e link al sito, per scaricare i risultati. Nel progetto Corsi on-line, il sistema fornisce soprattutto uno stimolo all'interazione diretta in classe (lo strumento è un supporto alla didattica in presenza). Il progetto FormAmbiente prevede un'interazione utenti-tutor da fare on-line ma anche da sviluppare nei momenti in presenza, durante i corsi brevi.
Concludo ritornando alla questione dei materiali. Non servono la spettacolarizzazione e gli effetti speciali; serve invece un'attenta taratura dei contenuti e dei tempi di erogazione, in relazione agli utenti e alle loro esigenze. Un sito altamente multimediale può ri-

[2] Si tratta dell'Osservatorio sulla Comunicazione dell'Università Cattolica di Milano.

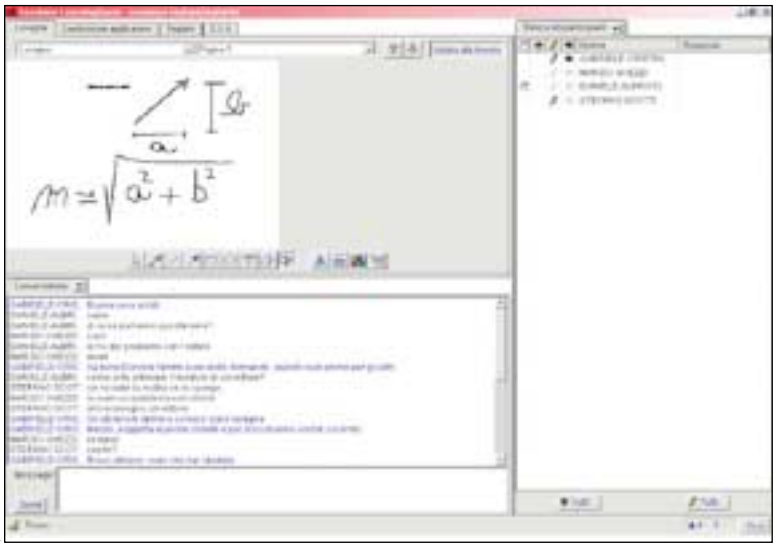
chiedere elevati tempi di accesso e di fruizione, magari incompatibili con un percorso formativo lungo, mirato a utenze che possono avere connessioni e macchine differenti. Meglio allora una scelta che veicola alcuni contenuti (quelli più assestati) attraverso CD-rom e altri (quelli maggiormente dinamici) via rete.

## 4. A COSA CI SERVE LA TECNOLOGIA (E QUANTO CI COSTA)

Che il computer di ciascuno degli attori di un processo di web learning sia equipaggiato con dei software "user friendly" lo do per scontato. Il punto è un altro: come direbbe Norman, si tratta di "rendere più vicino il giorno in cui la tecnologia informatica scomparirà dalla vista e la nuova tecnologia che ne prenderà il posto sarà immediatamente assimilabile e facile da usare ..." [13]. In attesa di questo momento (ma Stoll dubita che ci sarà mai), si tratta di piegare la tecnologia alle esigenze degli utenti. A questo punto il nostro obiettivo è abbastanza facile da enunciare: la tecnologia deve fornire gli strumenti (dalla banda larga alla tavoletta grafica) per favorire l'interazione.
Procedendo anche qui in maniera un po' schematica, ecco le principali ***necessità***:

- interazione uno-a-uno (studente-docente);
- interazione di gruppo (tra gli studenti);
- accesso alle informazioni e scaricamento dei materiali;
- monitoraggio e tracciamento (del singolo e della classe).

**FIGURA 3**
*La piattaforma Learning Space durante una sessione "live"*



A queste esigenze corrispondono alcuni strumenti ben noti dell'on-line learning, più o meno presenti in tutte le piattaforme [3]. Essi sono:

- la ***condivisione della lavagna*** (i partecipanti alla sessione di lavoro devono poter intervenire reciprocamente; è presente un livello superiore di gestione);
- la ***chat***, sia testuale che audio (almeno in broadcast), mentre è meno importante quella video;
- la ***bacheca***, con una duplice funzione (affissione di messaggi in uscita e deposito di testi in arrivo);
- il ***forum***, come luogo di dibattito "asincrono" (sia un forum generale che un forum per ogni materia);
- un sistema di ***test in tempo reale*** (con la possibilità di vedere immediatamente risultati e statistiche);
- un sistema di ***mail dedicate*** al progetto (indipendenti dalle mail personali degli utenti);
- la ***web-conference***, possibilmente integrata con il sistema di messaggistica della piattaforma (al fine di permettere domande e risposte in tempo reale);
- la definizione di differenti "***profili di utenza***", con la possibilità di collegarli a un database nel quale convergano informazioni quali-quantitative sul singolo utente;
- un sistema semplice e rapido di ***reportistica***, sul singolo, sulla specifica prova, sulla classe (è in pratica una matrice tridimensionale).

Altri servizi, come quello di aggiornamento rapido della propria pagina web o tutti quelli di gestione della classe, possono ulteriormente favorire il processo: ma una piattaforma che fornisca in modo affidabile i servizi sopra elencati può bastare.
In figura 3 è mostrata l'interfaccia della piattaforma Learning Space utilizzata nelle sessioni "live" della LLP: sono visibili alcuni degli strumenti indicati (la lavagna condivisa, la chat, il test in tempo reale, la gestione degli studenti presenti).
Quali sono i costi di un progetto di web learning ? Per rispondere a questa domanda è necessario distinguere tra costi di investimento (in hardware e software), costi di produzione dei materiali, costi di erogazione. Esistono poi, essenzialmente per i docenti, dei costi non monetizzabili.
Il dettaglio dei costi monetari dipende fortemente dalla dimensione, dal target di utenza e dalla copertura che si vuole dare al proget-

to. Come termine di paragone si può prendere la Cardean University americana [4], creata da un pool di università e di imprese dell'area ICT e diretta da D. Norman: nella sua realizzazione sono stati investiti circa 50 milioni di euro. Il progetto LLP ha richiesto un investimento di circa 5 milioni di euro (in larga misura sostenuti da Somedia, partner del Politecnico). I costi nella fase di avvio sono elevati: il recupero degli investimenti non può, quindi, che richiedere tempi lunghi e un mercato in crescita.
Anche per quanto riguarda i tempi di produzione la variabilità è elevata, tuttavia qualche indicazione può essere fornita. Il rapporto fruizione/preparazione va da 1:30 a 1:100 (per un'ora di fruizione possono essere necessarie da 30 a 100 ore di preparazione), con una forte dipendenza dalla modalità/rigidità del processo produttivo adottato. In ogni caso, per il docente il rapporto non è mai meno di 1:10. Progetti didattici del tipo videocassette (il caso 4 di figura 1) sono più veloci: per esempio, il Consorzio Nettuno dichiara un rapporto tra 1:20 e 1:30.
Un tempo di preparazione di questa rilevanza richiede ovviamente un team formato dal docente e da un gruppo di collaboratori: nei vari casi a cui ho lavorato il team era formato in media da 3-5 persone, con compiti diversificati (registrazione dell'audio, costruzione del glossario, analisi dei siti interessanti, creazione degli esercizi, ecc.).
Per i vari motivi indicati, lo sforzo organizzativo non è mai modesto. La gestione del processo produttivo è quindi un punto delicato. Per il team dei docenti contano assai di più le motivazioni didattiche e le condizioni di lavoro che non la remunerazione economica: qualunque intervento didattico "in presenza" avrebbe un rapporto tra ore di preparazione e retribuzione molto più favorevole.
Un discorso a parte va fatto per i costi non monetizzabili: ne accenno qui utilizzando un modello della Teoria dei giochi noto come "Il dilemma del prigioniero", presentato per esempio in [2], che in questo contesto si potrebbe ribattezzare "Il dilemma del (giovane) docente".
Lo schema del modello riporta una matrice dei benefici di due giocatori, ciascuno dei quali ha davanti a se la scelta (da ripetere ogni giorno) tra due alternative: cooperare con il suo antagonista o defezionare, cioè fare i propri escusivi interessi. I due non possono comunicare tra loro. Il sistema giudiziario premia il pentitismo, cioè la defezione.
Se i due cooperano i benefici per entrambi sono misurabili in 3 unità (in una scala tra 0 e 5). Se uno dei due sceglie un atteggiamento cooperativo e l'altro no, colui che defeziona ottiene il massimo beneficio (5) mentre l'altro viene duramente punito (beneficio 0). Se infine entrambi defezionano i benefici sono molto minori (il risultato è di 1 per entrambi): ma si tratta dell'unico punto di equilibrio. La strategia più logica per ognuno dei due è quindi quella di defezionare (questa scelta domina la scelta di cooperare), ma la situazione di mutua cooperazione sarebbe più favorevole per entrambi: se i due si fidano l'uno dell'altro la attueranno, se no avranno un beneficio minore (1 invece di 3).
In figura 4 e nel caso che voglio descrivere, i giocatori sono due (giovani) docenti, il sistema è quello della promozione universitaria, cooperare tra loro vorrebbe dire che entrambi adottano una didattica innovativa, defezionare vuol dire muoversi secondo le regole del sistema universitario (che premia quasi solo la ricerca) cioè minimizzare l'impegno didattico attuando una didattica tradizionale. La coppia di scelte dei due giocatori produce uno dei quattro risultati indicati nella parte (a) di figura 4.
Cosa c'entra tutto ciò con i costi di un progetto di web learning ? E perchè il modello vale per un "giovane" docente ? Nel nostro caso, cooperare significa "spendersi" nel progetto di web learning, defezionare significa adottare una didattica classica (meno dispendiosa in termini di tempo). Se il suo vicino di stanza non fa altrettanto, al giovane docente non conviene cooperare perchè impegnandosi in questa direzione scriverà meno articoli scientifici e farà meno carriera (se è giovane questo conta, se è già arrivato all'apice il modello ha meno significato). Se però entrambi decidessero che vale la pena di lavorare a un simile progetto, i benefici per la didattica di entrambi aumenterebbero e nel contempo non ci sarebbero danni in termini di carriera. Se poi qualcuno per loro (cioè sopra di loro) decidesse per la cooperazione reciproca il successo sarebbe garantito. Questo è il ruolo che l'istituzione universitaria può avere per far decollare progetti di web learning: modificare il quadro (passando

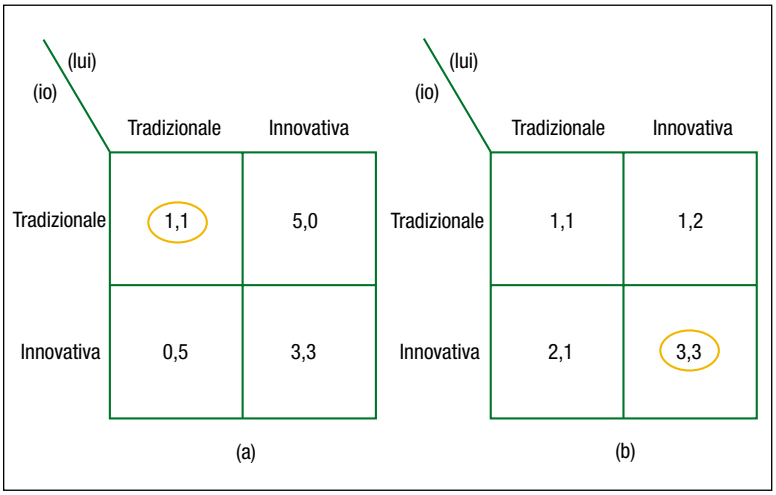



**FIGURA 4**
*La matrice di payoff per il dilemma del giovane docente*

cioè a una situazione come quella descritta nella parte (b) della figura), in modo da garantire visibilità e benefici immediati ai progetti di didattica innovativa.

## 5. EURO E NON SOLO

Cosa c'è in giro per il mondo nel settore di cui parliamo ? Una doverosa premessa è che qualsiasi esame è destinato ad una rapida obsolescenza, a causa della dinamica con cui le cose si stanno muovendo in questo momento. In generale si possono individuare tre tipologie di erogazione [15, 19]:

- università a distanza, cioè istituzioni accademiche che offrono esclusivamente corsi a distanza, sia attraverso la rete telematica (on-line learning), sia seguendo le modalità più tradizionali della formazione per corrispondenza (FaD tradizionale);
- università "dual-mode", cioè università che offrono corsi sia in presenza che on-line, in modo da cercare di soddisfare le esigenze di due fasce di utenza distinte;
- università in presenza con corsi "misti", cioè università che offrono esclusivamente corsi in presenza, alcuni dei quali prevedono una parte del programma su web (la rete ha la funzione di integrare i contenuti trattati in aula).

Cominciamo analizzando la situazione italiana. Viene spesso ricordato il Nettuno [8], un Consorzio che comprende più di 30 università, la Rai ed altri partner, da anni impegnato in un progetto didattico basato sulla produzione di videocassette (attualmente ha un catalogo che comprende molte migliaia di ore di videolezioni e di esercitazioni) diffuse attraverso due canali satellitari della Rai. Il progetto Nettuno è un ottimo esempio per il caso 4 dello schema di figura 1. I meriti storici del Consorzio non possono che essere riconosciuti. Il modello didattico, invece, mi lascia perplesso a causa di alcune differenze non marginali rispetto a quanto ho descritto in queste pagine: la poca interazione (anche se ultimamente il Consorzio ha modificato il suo modello, spostandolo in parte su Internet) e una responsabilità abbastanza diluita (quasi sempre gli esaminatori degli studenti sono docenti diversi da quelli che hanno registrato le videocassette, spesso ne ignorano il contenuto) mi sembrano i limiti più grossi dell'approccio.

Altre esperienze on-line sono presenti sul fronte dei corsi di master. Il Politecnico di Milano ha avviato alla fine del 2001 il master [16] NBA *Net Business Administration*, in collaborazione con Sfera società del gruppo Enel. La società Profingest, collegata all'università di Bologna, offre un master on-line [17] anch'esso nel settore della business administration.

Un'iniziativa particolare fa capo al Consorzio ICoN [10], costituito da un gruppo di università allo scopo di diffondere la cultura italiana all'estero: il Consorzio sta mettendo in rete varie unità didattiche e si propone di gestire attraverso le ambasciate e i consolati un sistema di certificazione per utenti residenti in ogni parte del mondo.

Ci sono poi alcuni progetti speciali che stanno avviandosi con il concorso di gruppi misti di università e imprese: tra questi cito il Consorzio LabCon, promosso dalla CRUI (la Conferenza dei Rettori), che vede una decina di atenei affiancare una società del gruppo Telecom. Un'indagine esplorativa promossa dai Presidi delle Facoltà di Ingegneria, allo scopo di creare un network di università, ha consentito di censire la situazione di quelle che già operano nel settore dell'e-learning.

Esiste poi il mercato "consumer", interpretabile solo in termini molto aggregati: l'annuale rapporto [1] dell'ANEE (l'associazione che riunisce i principali editori multimediali italiani) fornisce qualche dato.

In Europa fanno testo i numerosi documenti che la UE dedica alla formazione mediante le tecnologie dell'ICT. Le università europee di maggiore tradizione in questo campo sono la

| | Tipologia dei corsi erogati | Paesi guida | Università |
|---|---|---|---|
| Università virtuali | On-line(oppure per corrispondenza) | Canada<br>USA<br>Spagna | http://www.athabascau.ca/ (Athabasca)<br>http://www.cardean.edu/ (Cardean)<br>http://www.uoc.es (Un. Oberta) |
| Università "dual mode" | Sia in presenza che on-line | Canada<br>Australia | http://www.open.uoguelph.ca/ (Guelph)<br>http://www.csu.edu.au/ (Charles Sturt) |

**TABELLA 1**
*Alcune esperienze straniere*

Open University inglese [14] e la UOC - Universitat Oberta de Catalunya [20]. In entrambi i casi si tratta di università "general purpose", che offrono corsi su uno spettro molto ampio a studenti di varia provenienza e background. Ci sono poi vari tentativi di mettere in rete atenei di diversi stati, favorendo l'integrazione: tra i progetti in cui il Centro METID è stato direttamente coinvolto cito Teleregion-SUN (tra le regioni autodefinitesi "i motori d'Europa"), Vimims (con la partecipazione di alcune università non-UE), Corel (che ci vede collaborare proprio con la UOC), Tuelip (una rete di università con forti esperienze nell'on-line, con il supporto di IBM). In tutti questi casi si tratta di mettere in rete materiali didattici fruibili da studenti di altre università, a volte con forme di scambio. Ed è proprio il principio di reciprocità che sta alla base di questi progetti che ne è anche il punto debole: gruppi abbastanza eterogenei di atenei, creati per ottenere finanziamenti UE, difficilmente riescono a garantire le caratteristiche di qualità necessarie in un progetto di e-learning. Da qui una certa difficoltà a far decollare queste operazioni, dotandole di adeguati obiettivi didattici.
Non è quindi facile fornire le linee generali di sviluppo dei progetti europei e mondiali. Il settore è certamente in fase di espansione, anche se non così brillante come qualche previsione aveva indicato solo un anno fa. Nel mondo ci sono comunque tre situazioni degne di nota: quella americana, quella canadese, quella australiana. In tabella 1 sono riportati alcuni indirizzi significativi, riferiti alla classificazione proposta all'inizio di questa sezione.

## 6. PENTIRSI?

Le conclusioni di un articolo pongono sempre il problema di riassumere in poche righe un ragionamento articolato [6]. Provo a farlo partendo da queste premesse:
- parliamo di istruzione universitaria o post-universitaria, in cui prevale il trasferimento di informazioni (piuttosto che l'attenzione agli aspetti educativi) unito alla fornitura di metodi e strumenti per affrontare la complessità del lavoro;
- parliamo di un percorso formativo (e non di un più o meno gradevole fai-da-te), con responsabili precisi, tappe intermedie, servizi da utilizzare, valutazione e monitoraggio, momenti di confronto e di dibattito, un preciso obiettivo finale;
- parliamo a persone motivate (da ragioni varie, di carriera, di status, di curiosità, ma in ogni caso non "precettate dai genitori") per le quali è o dovrebbe essere chiaro che il percorso non è facile né divertente.

Riparto dalle principali obiezioni alla formazione on-line, in particolare da alcune fatte da Stoll: per ciascuna di esse provo a indicare i modi perchè "la goccia di pioggia finisca nel bacino giusto".

1. ***Assenza di una dimensione sociale***. Le scuole per corrispondenza e la vecchia FaD hanno fallito perchè non hanno considerato le relazioni tra i soggetti (lasciando lo studente isolato) e perchè hanno concentrato lo sforzo sui materiali (senza fornire a fianco i servizi). Bisogna invece pensare e progettare un apprendimento sia interattivo che collaborativo: spero di aver esemplificato a sufficienza quest'esigenza nell'articolo. La tecnologia lo consente, è necessario progettare la cosa fin dall'inizio.
2. ***Prevalenza del virtuale e passività***. Questa obiezione ("meglio una camminata nel bosco che la simulazione a video della crescita di una pianta") è interessante, ma riguarda tutta l'università (e non solo la FaDoL) in quanto basata molto sul sapere e molto meno sul saper fare. È vero che la formazione on-line ha in se

alcuni "germi dannosi" dell'istruzione programmata[3], ma ciò avviene prevalentemente secondo alcune modalità (per esempio quella basata sui "learning objects") che ipotizzano una parcellizzazione del sapere da riassemblare secondo le circostanze, senza una guida responsabile del percorso. Un progetto fortemente basato sull'interattività può scongiurare il pericolo della passività: la lentezza dei collegamenti è un problema in via di risoluzione e la possibilità di graduare (on/off line, sincrono/asincrono) va sfruttata.

**3.** ***Percorsi facili e superficialità***. Questa obiezione è assolutamente fondata, ma solo per qualche caso "da pubblicità". Non si dovrebbe promettere una formazione facile, da ottenere nei ritagli di tempo, ma un percorso duro, controllato, garantito nel risultato. Eventualmente si possono attivare percorsi diversi, con tempi differenziati (part-time), ma sempre sotto una precisa e unica responsabilità. Studiare non è "divertente" e la formazione on-line non è l'edutainment.

**4.** ***Valutazione e standardizzazione***. La considerazione di chi dice che nella FaD si promuove la capacità dello studente di rispondere solo ai test e alle domande standard non è priva di fondamento. In realtà è una questione di tempo e di progetto didattico, non una scelta obbligata: se il docente è disponibile a correggere elaborati più articolati (pur di superare lo scoglio della scrittura delle formule) la questione non si pone. È comunque utile che nel "registro" del docente ci sia spazio per valutazioni anche qualitative, basate sul livello della partecipazione dello studente ai forum e sul tipo di richieste che gli fa pervenire. Inoltre la valutazione finale dovrebbe essere in presenza (nel caso della LLP è così), con la doppia funzione di garantire dell'identità di chi ha inviato gli elaborati durante il semestre e di consentire un confronto faccia-a-faccia con lo studente.

**5.** ***Aspetti commerciali***. La FaDoL ha messo in moto un giro di affari crescente (anche se forse non del livello previsto), in cui si sono inserite anche le università: ciò è vero e dipende dalle molte (troppe) istituzioni universitarie costrette a competere sull'unica risorsa veramente scarsa, gli studenti. In aggiunta a ciò, la formazione on-line toglie anche il "filtro territoriale" (l'università della propria città o regione) che forniva un livello garantito di domanda. La competizione comporta dei costi e una visione attenta anche agli aspetti commerciali. Quando una cosa è gratuita (e la maggior parte delle cose che si trovano in Internet lo sono) il suo valore economico è basso: il costo elevato della FaDoL può essere anche una garanzia della sua qualità (molte volte nel progetto LLP ci siamo detti "gli studenti pagano salato e quindi hanno diritto a ...").

**6.** ***Investimenti vs. risultati***. La formazione on-line ha costi elevati, non c'è dubbio[4]. Ma i costi erano elevati anche quando si costruivano i primi edifici scolastici invece di pagare un precettore a ogni alunno (e l'obsolescenza dei contenuti non sarà così forte come si teme). Inoltre la ricerca, perchè di questo si tratta, ha alcuni elementi di rischio: perchè ciò che non si rimprovera al CERN dovrebbe essere messo in conto alla FaDoL? Il punto vero è capire a quali condizioni e come si può fare della buona formazione on-line.

Un discorso più difficile è quello sugli sviluppi futuri: le previsioni di espansione a ritmo vertiginoso sono finite insieme a tutte le altre sulle sorti del mondo web. Quello della formazione on-line è però un settore che risponde a esigenze vere: un certo ottimismo non è solo rituale. C'è poi tutto il discorso sulle tecnologie e sulla integrazione.

Recentemente si parla molto di "convergenza", cioè dell'integrazione TV-computer, e dei futuri benefici che essa produrrà. Più che le visioni un po' inquietanti (il tizio che fa home-banking mentre guarda una soap-opera e attende una pizza ordinata on-line), mi sembra plausibile e interessante lo scenario che vede l'invio dei materiali su canali a larga banda (via cavo o via etere) e i ritorni verso il docente (feedback, test volanti, domande) via web.

---

[3] Negli anni Cinquanta B.F. Skinner fu il principale promotore di una modalità di apprendimento basata su domande e successivi avanzamenti lungo un percorso programmato, una sorta di primitivo ipertesto.

[4] Stoll ([18], p.86) cita il caso di un investimento di 100 milioni di $ per un progetto che ha visto solo 100 studenti arrivare alla conclusione.

Attenzione però: oltre alla banda dovrà essere larga anche la memoria, una mediateca personale ingombra.
Come valutare dunque l'efficacia di un sistema di web learning? cioè di un insieme di contenuti culturali, materiali forniti, servizi offerti, organizzazione?
Direi essenzialmente in tre modi:
***a*** rispetto alla situazione precedente (cioè guardando alle opportunità che offre);
***b*** rispetto alle attese (cioè dal punto di vista degli obiettivi dichiarati);
***c*** rispetto agli standard (cioè attraverso un confronto con esperienze positive).
Nel nostro caso direi che l'opportunità di ri-avvicinare alla istruzione universitaria una parte di persone che avevano fatto (o stanno per fare) scelte diverse è sicuramente positiva, che l'obiettivo di un percorso facile è sbagliato ma non lo è quello di una buona interazione con docenti e tutor (certamente ottenibile e più favorevole di molte situazioni in presenza), che sono ancora in corso le definizioni di standard adeguati (e questo è un buon motivo per esserci).
Tutto bene, dunque ? No, certamente. Ho ben più di un dubbio su tutta questa vicenda. Ma d'altra parte, anche senza andare a casi eclatanti come la bomba atomica o la clonazione, non ho mai visto una scoperta o una tecnologia (o più semplicemente un'opportunità) che sia stata rifiutata per ragioni etiche o comunque concettuali. Il massimo che si può fare è volgerla in opportunità utile e cercare di dominarne gli effetti indesiderati. Meglio quindi operare criticamente fin dall'inizio piuttosto che scrivere libri di pentimento 20 anni dopo.

## Ringraziamenti

Molte persone hanno contribuito alle mie analisi. Tra le tante, ringrazio in particolare coloro che lavorano al Centro METID: le esperienze e le riflessioni che insieme abbiamo maturato sono la base di questo articolo.

## Bibliografia

[1] ANEE: *Editoria, contenuti e servizi nell'economia digitale in Italia*. Milano, 2001.

[2] Axelrod R: *Giochi di reciprocità*. Feltrinelli, 1985.

[3] Biolghini D, Cengarle M(eds.): *Net-learning*. Etas, 2000 (contiene un'analisi delle principali piattaforme).

[4] Cardean University, http://www.cardean.edu

[5] Centro Metodi E Tecnologie Innovative per la Didattica, http://corsi.metid.polimi.it, Politecnico di Milano.

[6] Colorni A: *Così nasce un "cyberateneo"*. Il Sole 24ore, 28 settembre 2001.

[7] Colorni A, Della Vigna P, Negrini R: *Www.laureaon-line.it: the on-line bachelor programme in computer engineering at Politecnico di Milano*. ICSEE 2002, 2002.

[8] Consorzio Nettuno, http://www.nettuno.stm.it/nettuno/index.htm

[9] Formez, http://www.formambiente.org, progetto di formazione su temi ambientali per i funzionari regionali.

[10] ICoN – Italian Culture on the Net, http://www.italicon.it

[11] Istituto per le Tecnologie Didattiche del CNR, http://itd.ge.cnr.it

[12] Laurea on-line in Ingegneria Informatica, http://www.laureaonline.it, Politecnico di Milano.

[13] Norman D: *Il computer invisibile*. Apogeo, 2000.

[14] Open University, http://www.opern.ac.uk

[15] Padovani N: *La formazione a distanza di livello universitario: il primo caso italiano di laurea ondine*. Tesi di laurea, relatore prof. F. Colombo, Università Cattolica del Sacro Cuore, Milano, 2001.

[16] Poliedra, Sfera, http://www.poliedra.polimi.it/l_master.htm, Net Business Administration.

[17] Profingest, Executive MBA on-line, http://www.profingest.it

[18] Stoll C: *Confessioni di un eretico high-tech*. Garzanti, 2001.

[19] Trentin G: *Insegnare e apprendere in rete*. Zanichelli, 1998.

[20] Universitat Oberta de Catalunya, http://www.uoc.es

Alberto Colorni è professore di Ricerca Operativa e direttore del Centro METID (Metodi E Tecnologie Innovative per la Didattica) al Politecnico di Milano.
Le sue ricerche vanno dai modelli matematici di decisione, alle applicazioni all'ambiente e ai trasporti, ai processi di e-learning.
È autore di circa 200 pubblicazioni; ha diretto progetti di ricerca italiani ed europei.
Ha vinto il premio Philip Morris per la Ricerca Scientifica nel 1989 e il Premio Cenacolo per l'Editoria e l'Innovazione nel 2001.
E-mail: alberto.colorni@polimi.it

# PARADIGMA ERP E TRASFORMAZIONE DELL'IMPRESA

**Gianmario Motta**

Il software ERP (Enterprise Resource Planning) hanno trasformato il sistema informativo aziendale. L'articolo illustra in primo luogo quelle caratteristiche distintive del paradigma ERP, che ne hanno favorito il successo: unicità del dato, modularità, prescrittività. Successivamente, discute le trasformazioni che gli ERP hanno indotto nel funzionamento delle imprese ai livelli dei processi operativi, manageriali, interaziendali. Infine considera i benefici potenziali offerti dalla trasformazione e propone uno schema di misurazione del suo valore, basato sulla valutazione dei vantaggi d'efficienza e d'efficacia.

## 1. LA SUITE ERP

L'acronimo ERP (*Enterprise Resource Planning*) è stato coniato agli inizi degli anni Novanta dal Gartner Group per indicare una *suite* di moduli applicativi integrati che supportano l'intera gamma dei processi di un'impresa. Oggi, una *suite* completa comprende decine di moduli applicativi, che possono essere schematicamente classificati nei tre gruppi di moduli *core* settoriali, moduli *core* intersettoriali e moduli *extended*.
Uno schema generale della *suite* ERP è in figura 1, dove la *suite* ERP è rappresentata da uno schema a T. I moduli settoriali sono la gamba della T, in quanto rappresentano la verticalizzazione delle applicazioni in ogni singolo settore industriale. La barra è formata dai moduli intersettoriali, in quanto orizzontali rispetto ai settori industriali, e dai moduli extended, che integrano l'azione degli ERP verso i mondi dei clienti e dei fornitori. La figura riporta anche un sommario elenco dei titoli dei principali moduli, che sono illustrati qui di seguito.

### 1.1. Moduli core intersettoriali

I moduli *core* intersettoriali sono sostanzialmente invarianti rispetto ai singoli settori industriali; in generale, informatizzano le attività aziendali di supporto.
I moduli istituzionali, così detti in quanto riflettono la regolamentazione pubblica, servono le attività amministrative, come la contabilità civilistica, la contabilità gestionale e la finanza aziendale, e la gestione delle risorse umane, che include sia le procedure contabili delle paghe sia i processi di gestione e sviluppo del personale.
I sistemi direzionali comprendono una serie di moduli che servono i processi, appunto, di conduzione manageriale dell'impresa, come la pianificazione strategica, la programmazione ed il controllo del budget, l'analisi dei costi e, più in generale, il reporting aziendale.
Sono intersettoriali, in quanto trasversali ai settori industriali, anche i moduli applicativi che pianificano e controllano le attività dei progetti ed elaborano la contabilità degli investimenti.
Fra i moduli intersettoriali includiamo anche il

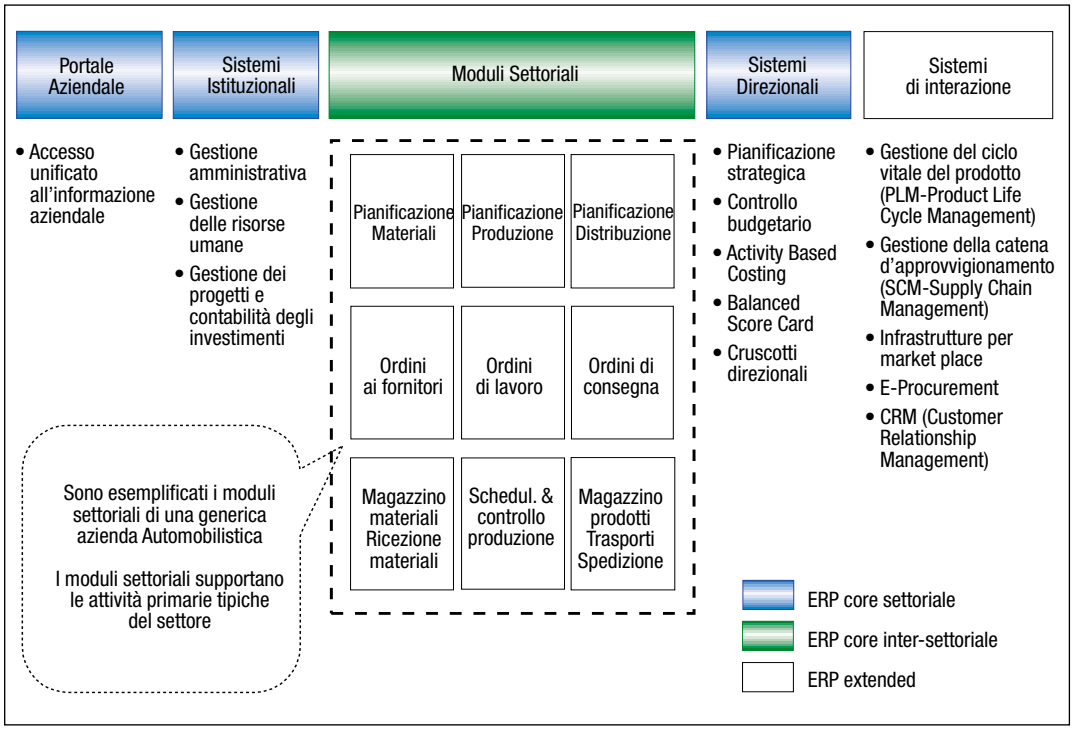



**FIGURA 1**
*La suite core ed extended*

portale aziendale. Apparso alla fine degli anni Novanta, offre un accesso via Web alle informazioni aziendali.
L'alta uniformità dei moduli intersettoriali ha favorito l'industrializzazione dell'offerta ERP in termini di qualità/costo; per contro, il cuore del sistema informativo aziendale, quindi il mercato maggiore, è dato dai moduli settoriali.

## 1.2. Moduli core settoriali

Le *suite* settoriali comprendono normalmente i moduli che supportano le attività primarie dell'azienda, tipiche del settore; sono molto diversificate, poichè riflettono le peculiarità d'ogni settore. Nella figura 1 è esemplificata, in sintesi estrema, la *suite* del settore automobilistico. Essa comprende una serie di moduli, che servono i processi d'approvvigionamento, produzione e vendita ai livelli di pianificazione, gestione degli ordini, attività fisiche.
La *suite* del settore automotoristico è diversa da quella del settore elettrico, che invece comprenderà moduli per la programmazione e controllo dei lavori (allacciamenti, nuovi impianti, dismissioni ecc.) e per la bollettazione.
Le *suite* settoriali sono numerose (per esempio il *vendor* SAP ne elenca una ventina) ed è conseguentemente elevato il costo sostenuto dai *vendor* per concepirle, realizzarle e mantenerle nel tempo. Non sorprendentemente, l'effettiva completezza delle *suite* settoriali è piuttosto variabile. In generale, è massima nel settore automotoristico, terra d'origine degli ERP, mentre è più limitata in altri settori, come banche o pubblica amministrazione, dove le *suite* includono ancora un limitato numero di moduli.

## 1.3. Extended ERP

L'*extended* ERP è formato da una serie di moduli che gestiscono le transazioni interaziendali e, più in generale, l'interazione fra più aziende o fra una singola azienda e clienti o fornitori.
In generale, queste *suite* supportano il ciclo vitale del prodotto (PLM, *Product Lifecycle Management*), la catena d'approvvigionamento (nota come SCM, *Supply Chain Management*), le interazioni con il cliente (CRM, *Customer Relationship Management*), l'*E-Procurement* e forniscono infrastrutture informatiche ai cosiddetti *Market Place*. Queste suite sono apparse sul mercato a partire dal 1995 come applicazioni indipendenti e separate dagli ERP core; con gli anni 2000, sono state integrate.
Valore distintivo dello schema ERP *extended* è l'integrazione fra transazioni interaziendali e transazioni interne. Per esempio, le *suite* ERP

| Numero di installazioni (I) o clienti (C) | Vendor | | |
|---|---|---|---|
| | Oracle (C) | People soft (C) | SAP (I) |
| **1999** | | | |
| WW | (ND) | 4.000 | 20.415 |
| Europa | (ND) | 900 | 13.500* |
| Italia | (ND) | (ND) | 517 |
| **2000** | | | |
| WW | 10.000 | 4.700 | 28.968 |
| Europa | 3.600 | 1.000* | 19.300* |
| Italia | 180 | 40 | 807 |
| **2001** | | | |
| WW | (ND) | 5.700* | 38.251 |
| Europa | (ND) | 1.300* | 25.500* |
| Italia | (ND) | 50 | 1.256 |

* Stime basate sulle dichiarazioni dei vendor

**TABELLA 1A**
*Diffusione dei maggiori vendor ERP*

| Vendor | Vendite (Milioni $) | Quota mercato |
|---|---|---|
| SAP | 5.839 | 30% |
| Oracle | 2.870 | 15% |
| Peoplesoft | 1.736 | 9% |
| J.D.Edwards | 980 | 5% |
| Altri | 7.127 | 41% |

**TABELLA 1B**
*Vendite 2000 dei maggiori vendor ERP (da AMR 2001)*

*core* sono integrate con i moduli CRM, che gestiscono i canali di contatto con il cliente (*call-center*, internet, agenti, negozi). L'integrazione fra CRM ed ERP *core* assicura l'effettiva esecuzione delle richieste del cliente (p.e. l'ordine di un'automobile, raccolto dai sistemi CRM, è programmato e controllato dai sistemi ERP *core* che gestiscono la produzione e la distribuzione). In altri termini, i sistemi CRM e, in generale, i sistemi d'interazione formano il *front-end* della azienda verso clienti e fornitori, mentre i sistemi ERP *core* formano il *back-end*.

## 1.4. Diffusione

Con l'estensione dello schema ERP, le aziende quindi hanno a disposizione una gamma molto ampia d'applicazioni informatiche; infatti:

- le *suite core* informatizzano le attività aziendali interne di livello operativo e direzionale;
- le *suite extended* informatizzano le transazioni interaziendali verso fornitori e clienti.

Non sorprendentemente, gli ERP sono divenuti uno dei massimi mercati d'applicazioni software. Alla fine degli anni Novanta, nel mondo si spendevano annualmente oltre 23 miliardi di dollari e i sistemi ERP erano installati in oltre 30.000 imprese. Oltre il 50% delle aziende europee ha uno o più moduli ERP e oltre il 35% li usa in tre o più aree funzionali [11]. Il numero totale di *vendor* sul mercato può essere stimato prossimo al centinaio. Tuttavia, solo alcuni, fra cui SAP, Oracle, Peoplesoft e JDEdwards, sono in grado di offrire l'intera gamma ERP. Il loro scarso numero riflette gli enormi investimenti necessari allo sviluppo ed al mantenimento di una gamma completa di moduli, intersettoriali e settoriali (Tabella 1).
Terza azienda software del mondo, SAP è nata intorno all'idea ERP e ha una quota stimata intorno ad un terzo del mercato. Il primo prodotto SAP, R/1, è stato un software *batch* per *mainframe*, sostituito nel 1981 da R/2 *on-line* e, nel 1992, da R/3 *on-line* e su architettura client-server, che negli anni 2000 è stata integrata da architetture *web-enabled*. A fine 2001, SAP, leader di mercato, dichiarava 12 milioni d'utenti, 44.500 installazioni, 1.000 aziende in partnership di vario tipo, 21 suite settoriali.
Oracle lancia la sua *suite* ERP negli anni Novanta, iniziando anch'essa dal core ERP poi integrato a portali e CRM, sfruttando ampiamente le proprie tecnologie di basi dati e In-

ternet. Punto di partenza di Peoplesoft sono i sistemi di gestione delle risorse umane, cui si aggiunge, dalla metà degli anni Novanta, la filiera standard ERP e CRM. JD Edwards, fondata nel 1977, si rivolge più specificatamente alla media impresa con un'offerta piuttosto ampia (6.000 clienti, di cui 300 in Italia).
Ai *vendor* multinazionali a gamma completa, si affiancano qualche decina di *vendor* multinazionali focalizzati su gruppi di moduli o su estensioni dello ERP *core*. Per esempio, Siebel, *leader* nel CRM, ha sviluppato un'offerta completa di CRM, che è a sua volta verticalizzata in soluzioni settoriali ed integrata con suite ERP. A sua volta, SAS, azienda *leader* di software statistici, offre una suite di CRM analitico, che elabora analisi sul comportamento del cliente o supporta la pianificazione delle campagne di promozione.
I grandi *vendor* multinazionali dominano nella grande impresa. Nelle piccole e medie imprese, invece, i primi 5 *vendor* hanno solo una quota minoritaria del mercato.

## 2. IL PARADIGMA ERP

La *suite* ERP rispecchia una precisa concezione del sistema informativo aziendale, con caratteristiche distintive:
- unicità dell'informazione;
- estensione e modularità funzionale;
- prescrittività.

Queste caratteristiche formano un paradigma funzionale, che indichiamo come paradigma ERP (un analogo concetto di paradigma ERP è discusso in [11]).

### 2.1. Unicità dell'informazione

Gli ERP sono caratterizzati da una base dati unica. Unica fisicamente od unificata attraverso un comune *repository* dei dati e servizi di replica automatica, memorizza i dati condivisi intorno alla quale ruotano i moduli. La base dati unica è una conquista sostanziale degli ERP che ha molti ed importanti vantaggi (Figura 2).
In primo luogo, l'aggiornamento unificato delle basi dati abilita la sincronizzazione di processi gestionali interdipendenti: p.e. l'arrivo di un materiale al magazzino aggiorna la situazione delle scorte, degli ordini ai fornitori e della contabilità dei fornitori, dando ai corrispondenti processi un'informazione unica e sincronizzata. Ciò non è possibile nelle tradizionali architetture ad isole, dove le basi dati sono separate e i dati comuni sono sincronizzati attraverso periodici processi d'allineamento. Nell'architettura ad isole, le informazioni sullo stesso oggetto (cliente, fornitore o materiale) sono temporalmente sfasate e ridondanti: il mancato pagamento di un cliente può non essere notificato in tempo alla gestione degli ordini, così che la situazione del cliente alla gestione degli ordini contrasta con quella della contabilità clienti (Figura 3).

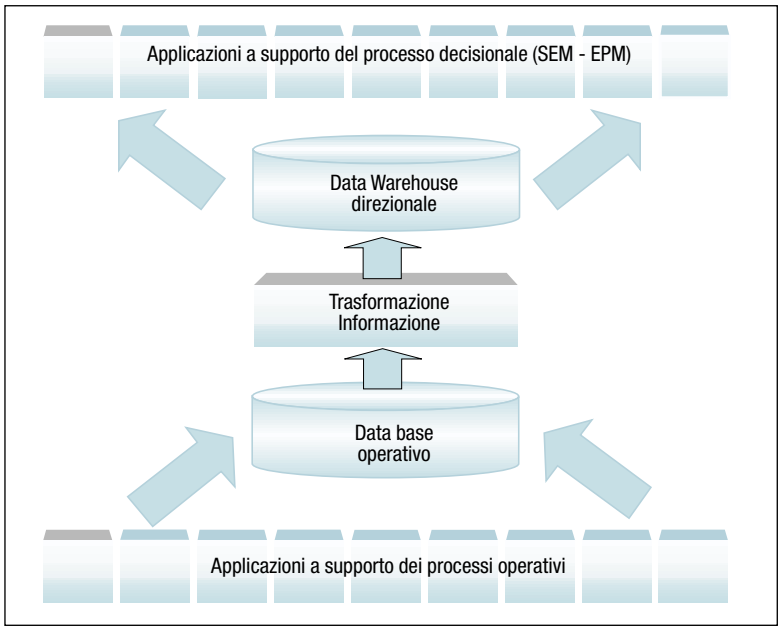



**FIGURA 2**
*Architettura ERP dei dati*

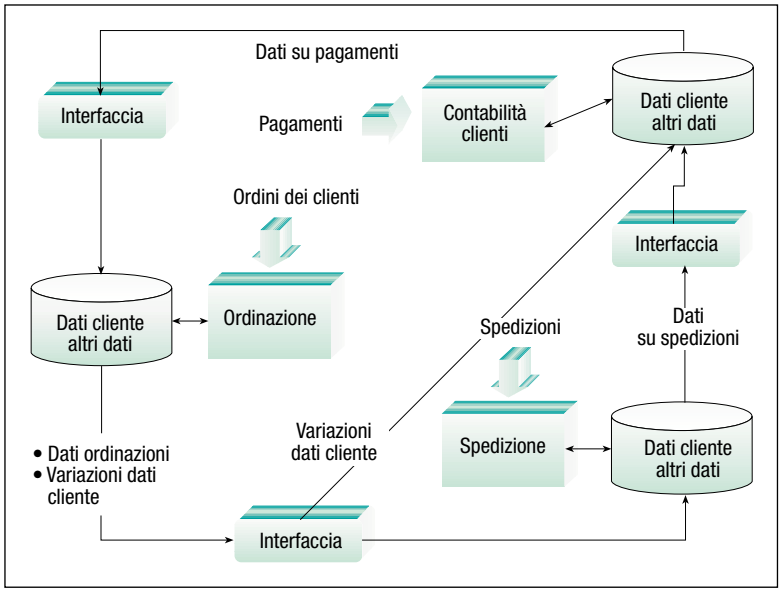



**FIGURA 3**
*Architettura ad isole di sistemi legacy: le basi dati sono sincronizzate da interfacce dedicate che propagano le modifiche*

In secondo luogo, l'architettura ERP certifica l'informazione e ne garantisce la tracciabilità: ogni evento di un processo per esempio la gestione di un magazzino, è testimoniato da un documento, per esempio una bolla di prelievo, che è specificatamente registrato nella base dati; ogni evento gestionale si riflette in una variazione di stato della base dati e la variazione é certificata da un documento.
Infine, l'unicità della base dati a livello operativo favorisce, in modo del tutto naturale, l'unicità dei dati per la direzione aziendale. L'unicità è ottenuta attraverso l'integrazione verticale dell'informazione operativa e dell'informazione manageriale. L'integrazione verticale si basa su un *Data Warehouse*, che memorizza i dati, aggregati e trasformati, estratti dalla base dati operativa (e da altre fonti). I dati sono elaborati da una *suite* di applicazioni, dette SEM (*Strategic Enterprise Management*) od EPM (*Enterprise Performance Management*), che assistono il management nella formulazione della strategia, nel *budgeting*, nell'analisi dei risultati.
L'integrazione rende disponibili informazioni sintetiche univoche (in quanto basate su dati operativi univoci ed unici), con un vantaggio rilevante per il management. Come molti studi hanno notato, la qualità dei dati è fra i vantaggi più apprezzati delle soluzioni ERP.

## 2.2. Estensione e modularità

Grazie all'estensione molto ampia, la *suite* ERP si propone come soluzione di riferimento per il sistema informativo aziendale, nelle sue componenti intra-aziendale, operativa direzionale, ed inter-aziendale.
Tuttavia, l'estensione funzionale sarebbe vana se la *suite* non fosse composta da moduli autosufficienti. Grazie alla modularità, l'azienda può scegliere una strategia d'implementazione coerente con la situazione dei sistemi e con il grado di rischio che è in grado di sostenere.
Una diffusa strategia semplice ed a basso rischio è l'implementazione parziale: l'azienda, cioè, sceglie di realizzare un piccolo numero di moduli, che vanno a sostituire preesistenti sistemi *legacy*.
La strategia, più ambiziosa, di implementare un elevato numero di moduli può essere attuata in due varianti, *one stop shopping* e *best of the breed*. Nel primo caso, privilegiando linearità e semplicità, l'azienda usa i moduli di un solo *vendor*, mentre, nel secondo, mette insieme moduli di più *vendor*, alla ricerca della soluzione ottimale per ogni processo aziendale, p.e. scegliendo il *vendor* A per la gestione del personale ed il *vendor* B per la gestione amministrativa.
Osserviamo che, data la modularità e l'ampia estensione funzionale degli ERP, la progettazione diventa simile ad una specie di *LEGO*, in cui è critico l'incastro fra i diversi moduli ERP, magari di più fornitori, e fra i moduli ERP e gli eventuali moduli *legacy*. Infatti, vanno garantite l'unicità e la sincronizzazione delle informazioni, attraverso interfacce standard, API (*Application Programming Interface*) e software di workflow o d'integrazione.

## 2.3. Prescrittività

I moduli ERP incorporano, in misura significativa, una logica di processo gestionale. Per esempio, la transazione di ricevimento dei materiali a magazzino presuppone un ordine al fornitore: un materiale non entra in azienda se non è stato ordinato e non può essere ordinato se non è stato richiesto da un ente aziendale autorizzato. Il software quindi norma il comportamento dell'utente aziendale, ribaltando la tradizionale concezione secondo cui è il software che si deve adattare all'utente.
La prescrittività ha importanti implicazioni. In primo luogo, semplifica l'analisi dei requisiti. L'analista, infatti, non deve specificare tutto il processo gestionale e tutto il sistema, ma si concentra sulle differenze rispetto al modello standard: definisce il processo ottimale, lo incrocia con le funzioni del sistema e sceglie le funzionalità del sistema. La progettazione funzionale diventa quindi una sorta di attività "taglia ed incolla" su di un *menu* di opzioni predefinite.
In secondo luogo, la prescrittività favorisce la standardizzazione dei processi ed uniforma i comportamenti, un vantaggio rilevante per le aziende distribuite territorialmente e le multinazionali.
Infine, la prescrittività può favorire una razionalizzazione dei processi, facendo coincidere il progetto informatico ERP con un progetto di

razionalizzaione organizzativa, meglio noto sotto la sigla BPR (*Business Process Reengineering*).
Tuttavia, la prescrittività comporta anche una certa rigidità, che può rendere gli ERP incompatibili con le specificità dell'utente. Infatti, se la razionalizzazione organizzativa richiesta è vasta e profonda, il progetto implica un costoso e rischioso intervento sul tessuto organizzativo dell'impresa. L'intervento può risultare infattibile per i tempi, troppo stretti, per i contenuti dei processi, incompatibili con il sistema dei valori esistente nell'azienda, per il rischio della trasformazione, troppo ampia, o, infine, per la mancanza di un gruppo di lavoro di quantità e qualità adeguata o, equivalentemente, per limiti di budget. L'alternativa ad adattare l'azienda al sistema è quella, piuttosto costosa e di un certo rischio tecnico, di adattare il sistema all'azienda, riscrivendo e/o modificando i moduli software.

## 3. TRASFORMAZIONE DELLA IMPRESA

### 3.1. La griglia di trasformazione

La diffusione degli ERP è evidente dai numeri ed è evidente la trasformazione delle imprese. Una volta lente, con una produzione inflessibile ed approvvigionamenti rigidi, in due decenni sono diventate capaci di gestire ordini personalizzati, lungo tutto il ciclo dal cliente finale al fornitore. E questo con scorte molto minori e una produttività molto maggiore. In generale, le caratteristiche degli ERP hanno contribuito a una serie di trasformazioni e queste trasformazioni, a loro volta, hanno generato, in varia misura, alcuni vantaggi. Le trasformazioni rilevabili appaiono riguardare i processi gestionali a diversi livelli (Figura 4):

- processi aziendali di livello operativo;
- processi aziendali di livello manageriale;
- processi interaziendali;
- modello di business (la trasformazione però appare molto più sfumata e controversa).

Le prime ricerche sulle trasformazioni aziendali indotte dalla IT sono degli anni Ottanta, con il riconoscimento del duplice ruolo della IT come tecnologia di produzione e come tecnologia di coordinamento [16]. La capacità delle IT, degli ERP in particolare, di trasformare i processi gestionali è discussa nella vastissima letteratura BPR (*Business Process Reengineering*), che segue lo storico articolo di Hammer "*Don't automate, obliterate*" [9] e si può dire conclusa da un articolo di Davenport [5] dal significativo titolo "*Putting the enterprise into the enterprise system*".

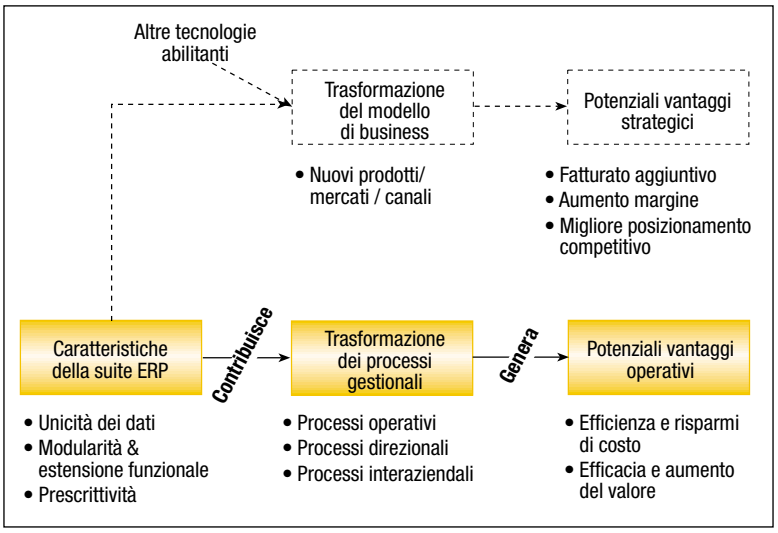



**FIGURA 4**
*La catena di trasformazione dei processi gestionali*

Vari autori hanno proposto schemi di correlazione fra caratteristiche di applicazioni IT e/o ERP e trasformazione dei processi. Venkatraman [21] ha ipotizzato una serie di stadi di *business transformation*, che vanno dalla ottimizzazione localizzata delle attività esecutive, all'integrazione fra attività interne, alla riconcezione dei processi, per arrivare alla riconcezione delle relazioni esterne e dello stesso business aziendale. Ciascuno stadio comporta una trasformazione più ampia e profonda dello stadio precedente e maggiori benefici potenziali.
Qui di seguito consideriamo le caratteristiche delle trasformazioni dei processi e dalla trasformazione del modello di business.

### 3.2. Trasformazione dei processi operativi

Con "trasformazione dei processi operativi" intendiamo un cambiamento dei processi che ne migliora l'efficienza e l'efficacia.
Grazie alla loro prescrittività, gli ERP dovrebbero cambiare l'organizzazione aziendale e portare ad un'organizzazione "processiva", più agile e flessibile, non parcellizzata, orientata a dare valore al cliente, con personale

| Variabile | Evoluzione |
|---|---|
| Flusso delle attività | • Minor numero di passi<br>• Minore durata del processo |
| Organizzazione operativa | • Arricchimento delle mansioni<br>• "Deparcelizzazione" del flusso del lavoro |
| Personale | • Versatilità operativa |
| Sistema di incentivazione e di controllo | • Obiettivi di servizio al cliente<br>• Obiettivi di performance sui processi gestionali |

**TABELLA 2**
*Evoluzione processiva delle variabili organizzative (adattato da Bracchi & Motta 2001)*

versatile e polivalente, in grado di svolgere un ampio ventaglio d'attività (Tabella 2).
I risultati dei progetti ERP indicano che l'evoluzione processiva esiste, ma parziale e condizionata da una serie di fattori; più precisamente:
❏ la trasformazione dei processi avviene e ne riguarda sia l'efficienza sia l'efficacia [6 - 13 - 18];
❏ un approccio sistemico al cambiamento, insieme con un iter di progettazione cauto, graduale, burocratico possono portare al successo, in quanto minimizzano il rischio [14];
❏ per ottenere veramente la trasformazione sono necessari un orientamento informatico appropriato del management e un ben orchestrato lavoro d'attori organizzativi che facilitino il cambiamento [2]; L'organizzazione e la gestione, nel senso più ampio, del progetto sono cruciali per il successo o il fallimento dei progetti ERP [14]: il presidio deve essere esteso a tutte le fasi dei progetti ERP, dal lancio all'accettazione ed alla stabilizzazione [18];
❏ un'elevata trasformazione dei processi, connessa all'adozione di ERP, aumenta il rischio del progetto, e richiede il presidio di un'ampia gamma di fattori di successo [19].
In sintesi, la trasformazione, per avvenire, richiede, in aggiunta al progetto informatico, uno specifico progetto organizzativo. Va osservato che l'attuazione di una trasformazione organizzativa significativa richiede l'impegno del management, non scontato e comunque costoso. Conseguentemente, è conveniente affrontare i progetti ERP, ad alto rischio organizzativo ed a medio rischio tecnico, riducendo i rischi, graduando i tempi, selezionando l'innovazione e coordinando, attraverso un'opportuna metodologia integrata, le filiere d'attività del progetto: implementativa, organizzativa, e infrastrutturale [3].

### 3.3. Trasformazione dei processi direzionali

Il contributo degli ERP alla trasformazione dei processi direzionali sta appunto nel loro contributo a rendere più efficiente e/o efficace l'informazione in input al processo decisionale e/o il processo decisionale stesso.
Un concetto rilevante per posizionare il contributo degli ERP è quello di IPC (*Information Processing Capacity*), che esprime l'adeguatezza di una organizzazione ad elaborare le informazioni richieste dai propri obiettivi e dal contesto in cui opera [3]. La capacità di un'azienda di operare in situazioni d'incertezza ambientale e di gestire strutture complesse è proporzionale alla sua IPC. In alternativa, l'azienda può investire in "risorse cuscinetto" (*slack resources*) come le scorte, che assorbono l'incertezza e diminuiscono il fabbisogno informativo, ma peggiorano le prestazioni d'efficienza e d'efficacia [8].
L'adozione degli ERP, specificatamente della suite "Sistemi Direzionali", aumenta la IPC attraverso l'aumento di valore della informazione, sotto vari aspetti:
❏ *ampiezza del dominio informativo*: l'ERP abbraccia tutta la catena del valore aziendale ed è integrabile con fonti esterne, pubbliche (Internet) e di fornitori;
❏ *disponibilità ed accessibilità*: l'informazione, memorizzata in basi dati unificate, è distribuibile in modo semplice, attraverso Internet ed accessi wireless;
❏ *velocità*: gli ERP accelerano il processo di creazione della informazione direzionale;
❏ *utilizzabilità*: gli ERP producono informazione fruibile anche per processi direzionali di livello strategico, come la Balanced Score Card [10].
La migliore qualità e, in definitiva, il minore costo dell'informazione offre un'opportunità anche per la trasformazione del processo decisionale. Secondo alcune teorie [12],

l'IT può ampliare il ruolo del management operativo (*empowerment*) ed accorciare le linee gerarchiche, introducendo un ciclo di controllo più breve e basato sulla collaborazione (Figura 5).

### 3.4. Trasformazione dei processi interaziendali

I processi interaziendali includono un arco ampio di relazioni [21] che spazia dalle transazioni interaziendali, alla gestione interaziendale di processi (scambi d'informazioni), a processi interaziendali (p.e. sviluppo congiunto di un nuovo prodotto) a sistemi di conoscenza condivisa.
Tale ambito di trasformazione riguarda primariamente i moduli *extended* ERP.
In primo luogo i moduli che gestiscono le transazioni con fornitori e partner permettono di realizzare sistemi inter-organizzativi per la cooperazione fra più aziende mediante condivisione d'informazioni (*e-procurement*), governo di flussi interaziendali di fornitura (EDI, sistemi di tracciamento d'ordini e di spedizioni), reti per svolgere un compito o servizio comune (pool di manutenzione impianti). Anche in questo caso, per ottenere una reale trasformazione, il progetto implementativo va integrato con un progetto logistico-organizzativo e con una valutazione della fattibilità del *business model* correlato [22].
In secondo luogo, i moduli CRM modificano sia il modo con cui l'azienda risponde al cliente sia il modo con cui il cliente interagisce con l'azienda. La trasformazione avviene attraverso l'attivazione di nuovi canali e delle corrispondenti tecnologie, come telefono (*call –center*), E-mail, Web fisso e mobile e, infine, sistemi per l'automazione delle forze di vendita, che, pur pre-esistenti, sono stati rinnovati tecnologicamente ed integrati con gli altri moduli CRM.
Un'analisi della trasformazione CRM esce, per la sua vastità, dagli scopi del nostro articolo. Ci limitiamo a sottolineare, che, poichè la risposta alle transazioni con i clienti e i fornitori è data dai processi interni, l'integrazione fra moduli *ERP core* e moduli di interazione è un requisito d'efficacia altrettanto importante della funzionalità dei moduli CRM, SCM che informatizzano il *front-end* rispettivamente verso il cliente od il fornitore.

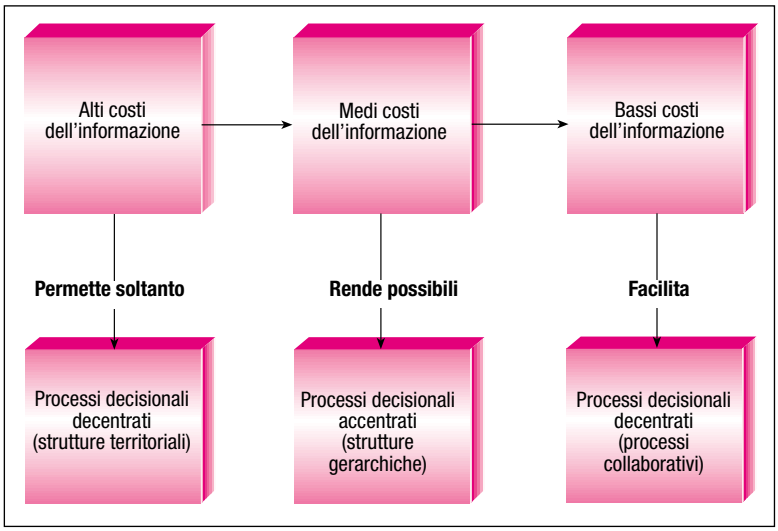



**FIGURA 5**
*Evoluzione dei processi decisionali nella teoria dell'empowerment di Malone*

### 3.5. Trasformazione del modello di business

Una trasformazione del modello di business implica la sostanziale innovazione del modello esistente. In sintesi, l'IT:
- rende possibili nuove attività con conseguenti nuovi ricavi;
- determina una domanda di nuovi prodotti e servizi;
- permette di sviluppare nuovi business a partire dai business esistenti.

Esempio significativo è Amazon, che innova la vendita per corrispondenza di libri e simili attraverso la tecnologia CRM. Un secondo esempio è la banca multicanale, che aggiunge alla tradizionale filiale i nuovi canali dei promotori finanziari, della banca telefonica (*call center*), della banca *on line* su *web*, sfruttando la capacità dei sistemi CRM di gestire in modo integrato il rapporto con il cliente. Un terzo esempio, emergente ora, è l'info-trattenimento: grazie alle nuove tecnologie di comunicazione (Fibre, ADSL) le aziende telefoniche offrono nuovi prodotti, come accesso Internet, intrattenimento e informazione. Con lo standard I-Mode, NTT introduce nella telefonia mobile il business dell'info-trattenimento, che raggiunge, a fine 2001, 30 milioni d'utenti.
Com'è evidente dagli esempi citati, la trasformazione è un'opportunità non di tutte le aziende ma solo di alcune, che sono connota-

te dall'elevata intensità informativa del prodotto e dei processi [16]. In secondo luogo, l'esistenza dell'opportunità non implica senz'altro il suo perseguimento: l'opportunità dell'e-business c'è per tutte le librerie, ma è stata perseguita soltanto da Amazon ed alcuni altri. Il perseguire un'opportunità riflette la strategia competitiva della singola azienda [2].
La trasformazione del modello di business passa attraverso i sistemi d'interazione con il cliente o con i fornitori. In questo caso, il contributo specifico degli *ERP core* è limitato. Tuttavia, un nuovo modello di business richiede anche (la progettazione di) un'efficiente catena del valore, che garantisca all'azienda la necessaria redditività. Tale requisito implica, a sua volta, una rivisitazione dei processi operativi interni ed una specifica strategia *ERP extended*. Amazon, che soltanto nel 2002 ha raggiunto il profitto, esemplifica quanto poco scontato sia il progetto operativo interno.

## 4. VANTAGGI E BENEFICI

La percezione comune associa ai progetti IT vantaggi d'efficienza e d'efficacia. Tuttavia, l'analisi aggregata dei dati di produttività d'aggregati nazionali od aziendali da' indicazioni contrastanti. Negli anni Ottanta, [20] evidenzia la non correlazione fra livello d'investimento IT e *performance* aziendale. Negli anni Novanta emerge la questione del *productivity paradox*, [4] che può essere così riassunto: gli investimenti IT salgono vertiginosamente mentre la produttività degli addetti (definita come quantità d'output prodotta per unità d'input) stagna o cresce marginalmente, come testimoniano le serie storiche USA e d'altri Paesi OCSE [7]. Il paradosso, tuttora irrisolto, é parzialmente spiegato dalla difficoltà di misurare in modo completo gli input e gli output: per esempio, l'investimento in Bancomat abbasserebbe la produttività delle banche se il loro costo fosse conteggiato nell'input ma il loro prodotto non fosse conteggiato nell'output (per esempio se la misura d'output fossero gli assegni per addetto).
Da parte nostra, ci sembra più coerente orientare la valutazione dei vantaggi sui processi, che sono lo scopo fondamentale degli ERP. Come già accennato, sono stati rilevati benefici specifici d'efficienza e d'efficacia, ottenuti con l'implementazione di sistemi ERP [6 - 13]. Tali benefici esprimono vantaggi operativi rilevanti su processi sia interni sia interaziendali (Tabella 3).
Una valutazione completa degli ERP deve quindi rispecchiare l'impatto complessivo sul business [17] ed includere anche significativi benefici intangibili, come l'accelerazione e la disponibilità dell'informazione manageriale [15].
Per quanto detto, riteniamo coerente para-

**TABELLA 3**
*Benefici ERP (rielaborato da [13])*

| Bene-fisico | ERP – core | ERP extended & SCM (Supply Chain Management) |
|---|---|---|
| Efficienza operativa | • Minori costi di staff (Business Process Reengineering - BPR)<br>• Minori scorte<br>• Minori costi logistici<br>• Minori costi di approvvigionamento<br>• Maggiore produttività e flessibilità | • Minori costi di staff (Business Process Reengineering - BPR)<br>• Minori scorte<br>• Minori costi logistici<br>• Minori costi di approvvigionamento<br>• Previsioni della domanda più affidabile |
| Efficacia operativa | • Maggiore tasso d'evasione ordini<br>• Migliorata capacità di risposta al cliente (*client responsiveness*) | • Maggiore tasso d'evasione ordini<br>• Migliorata capacità di risposta ai partner della catena di fornitura<br>• Minore *time-to-market* |
| IT | • Standardizzazione delle piattaforme IT | |
| Valore della Informazione | • Condivisione globale dell'informazione | |
| Altre | | • Creazione di nuove opportunità di mercato |

metrare la valutazione dei benefici sul valore della trasformazione dei processi. Il valore della trasformazione include il valore dei guadagni d'efficienza, ottenuti attraverso la migliore ingegneria dei processi, e dei guadagni d'efficacia, percepiti dai clienti, che comprendono il maggiore valore degli output (adottiamo le classiche definizioni Efficienza = Output effettivo / Input ed Efficacia = Output effettivo / Output atteso). In sintesi, il valore della trasformazione **T** può essere quantificato come la somma dei guadagni d'efficienza **E** e del guadagno d'efficacia **V** che è percepito dal cliente, cioè **T** = **E** + **V**. I benefici d'efficacia **V**, riflettono il guadagno di valore derivante da migliori prestazioni del processo, quali migliore livello di servizio e migliore qualità in uscita, e misurano la differenza fra il prezzo **P2** che il cliente è disposto pagare rispetto al prezzo **P1** ante-progetto, ovvero **V** = **P2** – **P1**. Si osservi che questo schema di valutazione della trasformazione è coerente con la lista di benefici citata in tabella 3; fanno eccezione le nuove opportunità di business, che richiedono un approccio specifico alla valutazione, a valle di un *business plan*.

## 5. PROSPETTIVE FUTURE

Dall'analisi dei *report* e delle presentazioni dei maggiori *vendor* ERP emergono una serie di linee evolutive generali [1 - 11].
Una prima ovvia linea d'evoluzione è tecnologica. Gli ERP, nati con architettura *client-server*, devono accedere ed essere accessibili via Internet e devono fornire portali di impresa. Questa evoluzione è inclusa nel paradigma standard degli ERP tracciato in figura 1.
Un'altra linea d'evoluzione è il completamento delle *suite* intersettoriali e settoriali, che permette ai *vendor* di allargare il mercato. Nell'ambito delle suite intersettoriali, lo sviluppo dei sistemi CRM, SCM e simili è ancora all'inizio; inoltre, sono da sviluppare i paradigmi gestionali di reale cooperazione interaziendale, non solo di scambio di dati. Nell'ambito delle soluzioni settoriali, molti settori sono ancora lontani da una *suite* completa, che comprenda tutti i sistemi d'elaborazione delle transazioni e raggiunga una completezza paragonabile alla *suite* per le aziende *automotive*.
Una terza linea d'evoluzione è l'integrazione. Pochi vendor riusciranno a sostenere il peso di una gamma completa ed integrata. Inoltre, molte aziende scelgono o sono costrette ad adottare soluzioni miste, in cui é necessario collegare *suite* ERP di fornitori diversi e sistemi ERP con sistemi *legacy*. Queste condizioni aprono un mercato ampio per i software d'integrazione delle applicazioni.
Un'ulteriore evoluzione è l'allargamento del mercato alle piccole e medie imprese attraverso *suite* semplificate e/o attraverso nuove modalità di fruizione. A questo scopo, alla modalità a progetto si affianca la modalità ASP (*Application Service Provider*), cui l'azienda accede via Web. L'azienda si limita a pagare l'uso di un software preconfezionato e verticalizzato, che risiede sui server di un centro servizi; in questo modo, può abbattere drammaticamente i costi d'utilizzo ed evitare i costi, i tempi ed i rischi dei progetti.

## 5. CONCLUSIONI

Il fenomeno ERP rispecchia la progressiva uniformazione del sistema informativo aziendale ed interaziendale in un paradigma completo ed integrato che ha trasformato, in varia misura, le aziende che lo hanno adottato.
La prima sostanziale trasformazione è in realtà la trasformazione del sistema informativo aziendale, che da collezione d'applicazioni diverse ed indipendenti diviene un'ordinata catena di montaggio e distribuzione dell'informazione.
La trasformazione del sistema informativo può favorire la trasformazione dei processi a livello operativo, direzionale, interaziendale. La trasformazione direzionale é un'area di potenziale alto successo, che integra bene modelli manageriali maturi con una tecnologia adeguata. La trasformazione interaziendale è ancora agli inizi. Il contributo degli ERP all'innovazione del modello di business appare limitato e incidentale.
Le trasformazioni qui discusse sono potenziali. La trasformazione effettiva è funzione dell'effettiva capacità dell'azienda di sfruttare le potenzialità ERP attraverso un'opportu-

na trasformazione del tessuto organizzativo ed un approccio cauto e ben bilanciato al progetto ERP.
Per misurare i benefici, potenziali ed effettivi degli ERP, è conveniente considerare il valore totale della trasformazione, concepito come la somma algebrica dei vari guadagni d'efficienza operativa e d'efficacia indotti dai progetti ERP.

## Bibliografia


[1] Austin RD, Cotteler MJ, Escalle CX: *Enterprise Resource Planning: Technology Note*. Harvard Business School, Publ. 1999. update nov 2001.

[2] Benjamin RI, Markus ML: The magic bullet theory in IT-enabled transformation. *Sloan Management Review*, 1997.

[3] Bracchi G, Francalanci C, Motta G: *Sistemi informativi e aziende in rete*. McGraw Hill, Milano, 2001.

[4] Brynjolfsson E, Hitt LM: Beyond the Productivity Paradox. *Communications of the ACM*, Vol. 41, n. 8, 1998, p. 49-55.

[5] Davenport TH: Putting the enterprise into the enterprise system. *Harvard Business Review*, Vol. 75, n. 4, 1998.

[6] Deloitte Consulting: ERP's second wave, European Research Presentation. *Deloitte Consulting*, 1999.

[7] Dewan S, Kraemer KL: International dimensions of productivity paradox. *Communications of the ACM*, Vol. 41, n. 8, 1998, p. 56-62.

[8] Galbraith JR: *Designing Complex Organizations*. Addison-Wesley Publishing Company Inc., 1973.

[9] Hammer M: Reengineering Work: Don't Automate, Obliterate. *Harvard Business Review*, Vol. 68, n. 4, 1990, p. 104-112.

[10] Kaplan RS, Norton DP: Using the Balanced Scorecard as a Strategic Management System. H*arvard Business Review*, Vol. 74, n. 1, 1996, p. 75-86.

[11] Mabert Vincent A, Soni Ashok, Venkatraman MA: Enterprise Resource Planning: common myths versus evolving reality. *Business Horizons*, May-June 2001, p. 69-75.

[12] Malone TW: Is empowerment just a fad? Control, management, and IT. *Sloan Management Review*, Vol. 38, n. 2, 1997.

[13] Ming-Ling C, Shaw WH: *Distinguishing the critical success factors between e-commerce, enterprise resource planning, and supply chain management*. Engineering Management Society, 2000, p. 596-601.

[14] Motwani J, Mircahandani D, Madan M, Gunasekaran A: Successful implementation of ERP projects: evidence from two case studies. *International Journal of Production Economics*, Vol. 75, 2002, p. 83-96.

[15] Murphy KE, Simon SJ: *Using cost benefit analysis for enterprise resource planning project evaluation: a case for including intangibles*. Proceedings of the 34th Annual Hawaii International Conference on System Sciences, 2001, p. 2955 -2965.

[16] Porter ME, Millar VE: How Information Gives You Competitive Advantage. *Harvard Business Review*, Vol. 63, n. 4, 1985, p. 149-161.

[17] Reneyi D, Sherwood-Smith M: Outcomes and benefit modeling from information systems investment. *The International Journal of Flexible Manufacturing Systems*, Vol. 13, 2001, p. 105-129.

[18] Ross JW, Vitale MR: The ERP revolution: surviving versus thriving. *Information Systems Frontiers*, Vol. 2, n. 2, 2000, p. 233- 241.

[19] Somers TM, Nelson K: *The impact of critical success factors across the stages of enterprise resource planning implementations*. 34th Hawaii International Conference on Systems Sciences, 2001.

[20] Strassmann PA: *The information pay-off*. The Free Press, 1985.

[21] Venkatraman N: IT enabled business transformation: from automation to business scope redefinition. *Sloan Management Review*, Vol. 35, n. 2, Winter 1994, p. 74-88.

[22] Volkoff O, Chan YE, Newson PEF: Leading the development and implementation of collaborative interorganizational systems. *Information and Management*, Vol. 35, 1999, p. 63-75.



GIANMARIO MOTTA Laureato in filosofia, è stato dirigente e partner in multinazionali del settore ITC, fra cui EDS e Deloitte Consulting. Studioso di sistemi informativi direzionali e delle strategie aziendali IT, è docente di sistemi informativi al Politecnico di Milano, ed è autore, con Giampio Bracchi, di numerosi testi.
E-mail: gianmario.motta@polimi.it

# CARENZA DI COMPETENZE: "SKILL SHORTAGE" NEL SETTORE ICT

**Renzo Provedel**

L'articolo propone una lettura ampia e critica del fenomeno della carenza di competenze ICT e risponde a due domande: esiste ancora, in questo quadro economico mondiale incerto, un fabbisogno così rilevante di competenze e di specialisti ICT, e se sì quali sono le possibili azioni concrete per affrontarlo? Sono valutati l'urgenza dell'intervento, la metodologia ottimale per far crescere le competenze ICT nelle organizzazioni e se le nuove professioni ICT siano lo strumento principale per innovare i business della "old economy".

## 1. SKILL SHORTAGE: C'È ANCORA ?

Negli ultimi due anni un argomento ha "tenuto banco" nel settore delle Tecnologie dell'Informazione e della Comunicazione (ICT): la carenza di specialisti e di competenze di ICT sul mercato, che ha costretto le aziende della domanda e dell'offerta ad inventarsi soluzioni "ponte" per affrontare la crescita tumultuosa dei bisogni e dei fatturati.
Le prime valutazioni nel maggio 2000 indicavano uno skill shortage per l'Italia di oltre 50.000 unità che avrebbero oltrepassato le 100.000 nell'anno 2001.
Questo fenomeno, battezzato "skill shortage", è apparso ufficialmente per la prima volta nel 1999 [5] più o meno in concomitanza col termine "new economy" che evocava una forte discontinuità nello sviluppo delle economie grazie alla tecnologia digitale.
Questo articolo vuole proporre una lettura ampia e critica dei diversi aspetti di questo fenomeno per cercare di dare una risposta alla doppia domanda che oggi emerge dai media e dalla pubblicistica di settore: esiste ancora, in questo quadro di economia mondiale quasi recessiva e soprattutto incerta, un fabbisogno così rilevante di competenze e di specialisti ICT e se sì quali sono le possibili azioni concrete per affrontarlo?
Viene sviluppata una tesi e fornita una metodologia, che non pretende di essere l'unica, ma che ha un interessante punto di forza, quello di essere stata sperimentata ed adottata in un grande Paese come la Gran Bretagna.

## 2. IL "POPOLO" DELLA NET ECONOMY: NASCE NELL'ANNO 2000

Per capire il fenomeno bisogna far parlare un pò i numeri ! Gli occupati nell'ICT sono stati valutati in Italia sino al 1999 a poco più di 441.000 (il 2,1% del totale Italia) in quanto si misurava solo l'occupazione nelle aziende dell'offerta. Nell'anno 2000 la popolazione viene riclassificata, nuovi settori si inseriscono, ed il "popolo" ICT diventa [8] di 1.313.000

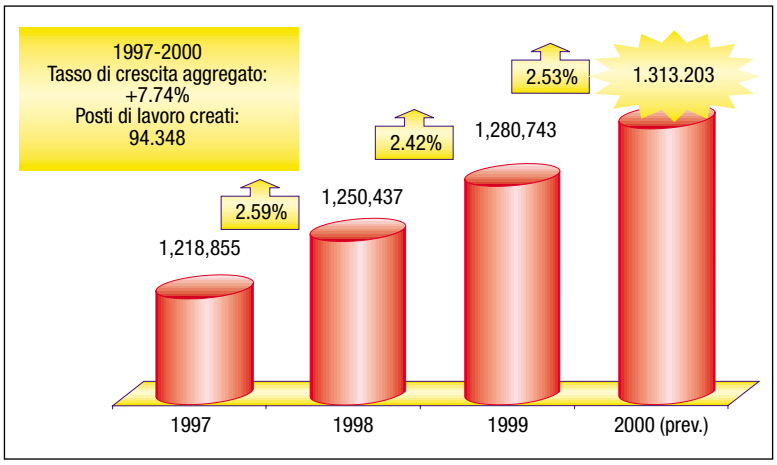



**FIGURA 1**
*L'occupazione nella Net Economy (1997-2000)*
*(Fonte: Federcomin/ NetConsulting su dati ISTAT)*

persone (il 6% degli occupati). Ma non solo: cambia anche nome e diventa *il popolo della Net Economy* (Figura 1).
Questo nuovo modo di vedere il settore fa emergere un dato molto importante: nel periodo 1997-2000 si sono creati oltre 94.000 nuovi posti di lavoro (+7,74%).
Merita capire come sia composta questa popolazione (Figura 2); le novità sono rappresentate dagli addetti dei fornitori di contenuti e servizi di Comunicazione (editoria, radio, televisione, pubblicità), dagli addetti "internet" nelle aziende della domanda, dagli addetti presso le "net company".
Quale futuro potrà avere ? È interessante leggere il rapporto *Employment Outlook 1998-2008 dello US Department of Labour* in cui si vede chiaramente come la Net economy non possa andare oltre il 10% del Pil nei paesi occidentali: in dieci anni l'Europa passerebbe dal 6% all'8% (e l'Italia dal 5% al 7%) e gli USA dall'8% al 10%.
I dati più aggiornati, forniti dall'Osservatorio

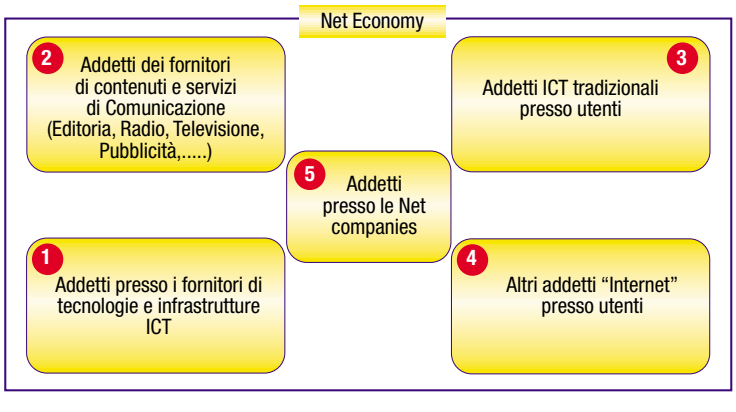



**FIGURA 2**
*L'occupazione nella Net Economy*
*Fonte: Federcomin/ NetConsulting su dati ISTAT*

Federcomin, nell'anno 2001 mostrano alcune novità (Figura 3 e 4).
La novità principale è l'occupazione "collegata" ad Internet ed alle nuove tecnologie che comincia a prendere forma e quantità: nell'anno 2001 viene misurata in almeno 178.000 unità sul totale di 1,4 milioni.

## 3. IL "GAP" DI COMPETENZE È EGUALE PER TUTTI?

La risposta è NO. Sulla base dei numeri, di cui sopra, si possono ipotizzare almeno due scenari per il "popolo" della net economy.
**1.** *Per un numero enorme di addetti, oltre 1,2 milioni*, il principale problema del futuro sarà la gestione del proprio sviluppo professionale per integrare ed acquisire nuove competenze tecniche nelle aree nuove, come Internet, le reti a larga banda, le reti wireless, le nuove modalità di servizio ASP, le opportunità applicative delle nuove tecnologie per lo sviluppo del business (e per il servizio ad imprese e cittadini per le diverse P.A.). Ciò sarà vero sia per la domanda sia per l'offerta pur con esigenze e percorsi diversi.
**2.** *Per gli specialisti* che già oggi operano più direttamente con le nuove tecnologie (almeno 200.000) le sfide sono almeno due: le evoluzioni rapidissime delle nuove tecnologie che comportano monitoraggio ed apprendimento continui, ed i nuovi modelli di business che bisogna "inventare" per sfruttare il potere abilitante delle nuove tecnologie.
Su questi due temi tornerò approfonditamente nel seguito per comporre e capire il quadro di insieme dello *skill shortage* e per identificare strategie adeguate a trattare i diversi fenomeni (vedasi paragrafo 6: "*Un caso emblematico: quale e quanta logistica per l'e-Business*").

## 4. GLI "ALTRI" HANNO PROBLEMI DI "SKILL GAP"?

Ovvero: chi deve affrontare il problema dello *skill gap* e come ? Quando uso il termine "altri" mi viene spontaneo pensare a "tutti gli altri" ....occupati e non occupati ! Può essere molto utile classificare gli "altri" in due grandi categorie:
***a.*** *la categoria degli occupati non "net eco-*

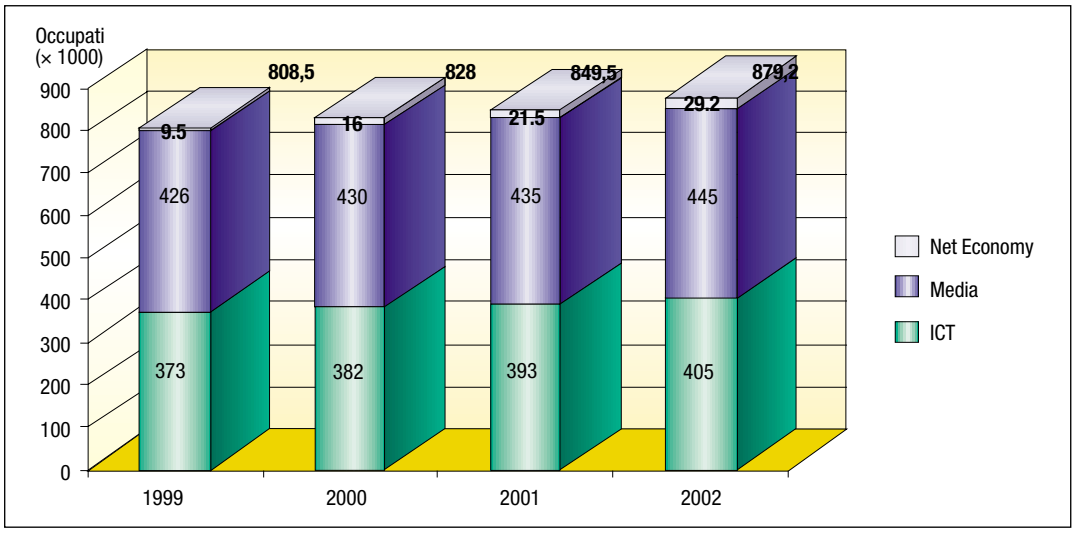



**FIGURA 3**

*La crescita dell'occupazione nelle aziende dell'offerta*

*(Fonte: Federcomin, 2001)*

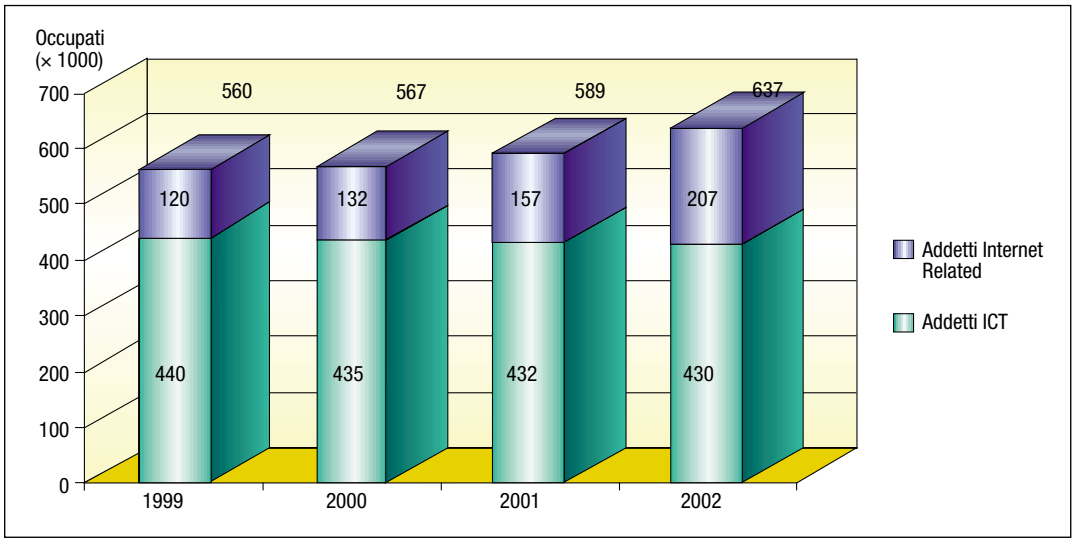



**FIGURA 4**

*La crescita dell'occupazione nelle aziende della domanda*

*(Fonte: Federcomin, 2001)*

*nomy*", che sono... oltre 19 milioni. Qualche volta si è usato il termine "old economy", ma in realtà bisognerebbe dire più semplicemente "economia" o economia tradizionale, non ancora "contaminata" dalle nuove tecnologie della net economy, come Internet e l'e-Business in generale.

***b.*** *la categoria dei consumatori*, cioè tutti noi. Oggi siamo sempre più convinti che lo sviluppo futuro dell'economia sarà spinto dalle nuove tecnologie solo se sapremo dare una risposta convincente su problemi fondamentali: la fiducia tra produttori e consumatori durante le transazioni in rete, la facilità d'uso delle nuove tecnologie, l'accesso generalizzato ed a basso costo per tutti i consumatori.

Lo "*skill gap*", termine che meglio rappresenta l'attuale situazione (più carenze di competenze che di numero di specialisti), riguarda entrambe le categorie ed è quindi un problema ed un fenomeno di massa per i prossimi anni. Tutti devono conoscere e saper usare le nuove tecnologie. *L'alfabetizzazione di base è una necessità vitale per l'intera popolazione*. Le competenze richieste sono di natura ed entità molto differenziate; molti programmi sono già in atto sotto l'egida di aziende, di enti della pubblica amministrazione, del governo centrale e regionale, di associazioni di categoria. C'è un fiorire di idee, dibattiti, iniziative. Anche la Unione Europea e l'OCDE (organizzazione di cooperazione e sviluppo economico) spingono decisamente in questa direzione. È stato coniato il termine "*digital divide*" per evocare pericolose arretratezze culturali che potrebbero mettere in un ghetto le persone

"incompetenti" e addirittura emarginare Paesi che non sappiano imparare la nuova cultura. C'è quindi una prima risposta al quesito iniziale dell'articolo:

> Lo skill gap ICT è una urgenza vera per l'economia italiana nella sua totalità e non solo per la net economy e va affrontata con diversi strumenti a tutto campo, distinguendo tra interventi per le imprese e per la P.A., e quindi per gli occupati, ed interventi sull'intera popolazione di consumatori.

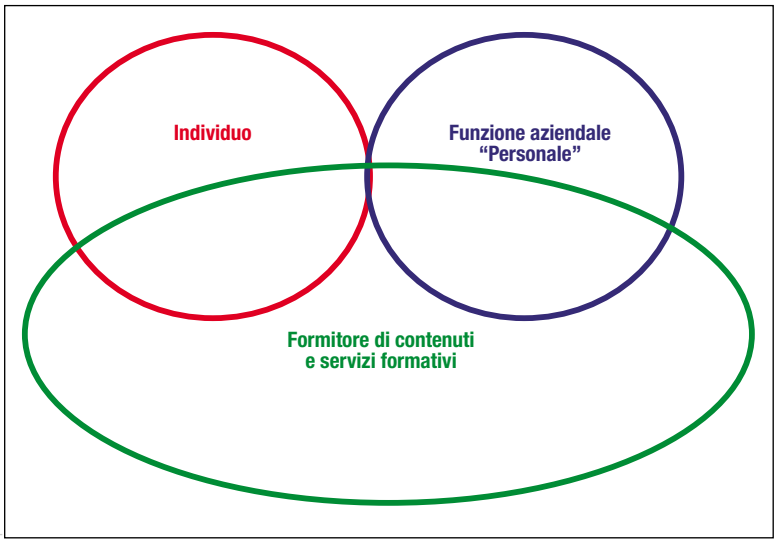
## 5. CON QUALI STRUMENTI SI MISURA E SI RIDUCE LA "CARENZA" DI COMPETENZE?

Ci dobbiamo porre una domanda: stiamo facendo una "ricerca" estesa a tutto il Paese o stiamo cercando una soluzione per la nostra organizzazione? I metodi devono essere diversi nei due casi.

### 5.1. Analisi Paese

Nelle ricerche estese ad un intero Paese il metodo con cui si cerca e si misura lo *skill shortage* è sintetico; si usano cioè campioni di aziende e di persone, e si usano questionari, che possono però rappresentare qualche volta il desiderio piuttosto che la necessità. Con la legge dei grandi numeri e con l'integrazione di altri metodi (Focus Group, Delfi, ad esempio) si converge verso un risultato affidabile. È il miglior metodo che si conosca oggi e quindi possiamo accettarne i risultati per fare diagnosi, e per capire i grandi fenomeni. Non tratteremo nell'articolo questo problema.

**FIGURA 5**
*Il modello. Sviluppo professionale e di carriera*
*(Fonte: Skillmatrix 2001)*




### 5.2. Analisi di una organizzazione

Quando però si deve passare all'azione si devono adottare altri metodi finalizzati ai singoli casi reali, per *confermare la diagnosi e fare la giusta terapia*.
Abbiamo diverse metodologie e casi esemplari di singole Aziende e di "filiere" produttive e di servizio. Da queste diverse situazioni di successo ho estratto tre "strumenti" che, a mio avviso, permettono di dominare il problema e trovare le soluzioni. In primis un modello di riferimento, poi gli standard professionali (utilizzati da oltre 15 anni in UK e da qualche anno in Italia), ed infine l'analisi funzionale.

#### 5.2.1. Un modello di riferimento per gli interventi nelle organizzazioni

Il modello di apprendimento più moderno si presenta come l'intersezione di tre domini (Figura 5):
**i** il dominio *individuale* delle esigenze legate al proprio ruolo, alle proprie competenze, al proprio sviluppo individuale;
**ii** il dominio *aziendale* in particolare quello funzionale del Personale, legato all'obiettivo di sviluppo e di miglioramento delle competenze organizzative;
**iii** il dominio dell'*offerta formativa*, in termini di contenuti e di modalità di apprendimento.
È stato sviluppato un modello dettagliato, per cui all'interno di ogni "circuito", si sono identificate una serie di attività. Un esempio completo per il circuito "individuo" si può vedere nella figura 6 (cortesia di Skillmatrix, 2001); esso mette in evidenza, ad esempio:
- il "profilatore di ruolo";
- il "valutatore a 360 gradi";
- il "Pianificatore di professionalità".

C'è ora una seconda risposta al quesito sul "come" si affronta lo skill shortage:

> Le organizzazioni dovrebbero adottare, per lo sviluppo delle proprie risorse umane, un modello di riferimento basato sull'apprendimento e sulla creazione di conoscenze che integri esigenze ed obiettivi individuali con esigenze ed obiettivi dell'intera organizzazione. Questo modello facilita gli interventi per la riduzione dello "*skill gap*".

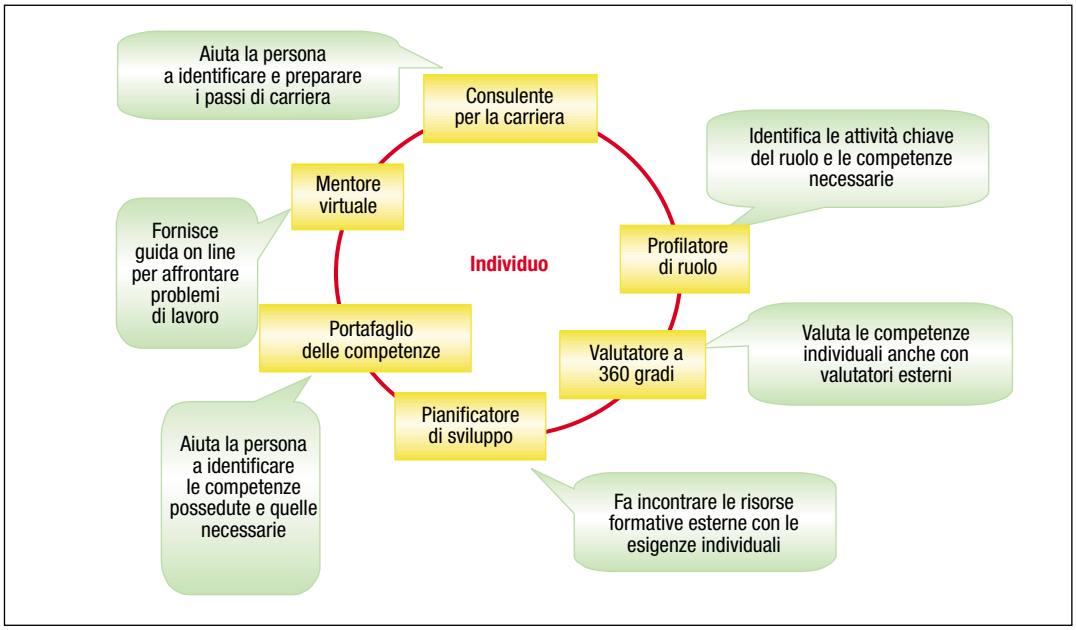



**FIGURA 6**
*Ciclo "individuo"*
*(Fonte: Skillmatrix 2001)*

### 5.2.2. Gli standard professionali

La metodologia più interessante per l'identificazione e lo sviluppo delle competenze, ed anche più sperimentata, è quella degli "standard professionali" [11].
È molto interessante ripercorrere il caso UK. Negli anni '80 in Gran Bretagna il governo, i datori di lavoro ed i sindacati si accordano per creare gli Standard Professionali Nazionali (*National Occupational Standards*): essi definiscono lo standard di prestazione richiesto ai lavoratori a tutti i livelli e per tutti i settori dell'economia.
Detto in altro modo questi standard definiscono le attese di ruolo, e quindi le competenze e conoscenze necessarie per lo svolgimento di ogni attività manageriale e professionale.
In uno studio su di una popolazione di più di 3.000 manager esemplari, la ricerca ha definito l'"obiettivo principale" di tutti i manager: *Raggiungere gli obiettivi della organizzazione e migliorare continuamente i suoi risultati*.
Usando il metodo di *Functional Analysis*, i manager hanno risposto alla domanda ripetuta a vari livelli: *Che cosa bisogna fare per ottenere l'obiettivo principale*? L'analisi ha riconosciuto e analizzato sette funzioni manageriali chiave: gestire attività, gestire risorse, gestire informazioni, gestire persone, gestire ambiente ed energia, gestire la qualità, gestire progetti. Per ognuna delle sette **Funzioni** fondamentali sono stati identificati complessivamente 66 **Processi chiave**. Ad ogni Processo Chiave sono collegati specifici **Sottoprocessi**, **Standard di performance per ogni sottoprocesso, Competenze "soft", Conoscenze ed abilità**.

**UN ESEMPIO**

**Funzione**: *Gestire attività*

**Processo chiave**: Attività per soddisfare il Cliente

**Sottoprocessi**: Concordare i requisiti con i clienti, Pianificare attività per soddisfare i bisogni del cliente, Assicurare la rispondenza di prodotti/servizi ai requisiti

**Standard di performance**: Assicurare supporto alle persone chiave per l'erogazione, Comunicare chiaramente e prontamente con i clienti, Intervenire prontamente di fronte a non conformità

**Competenze "soft"**: Comunicazione, Gestione di team, Influenza sugli altri, Focalizzazione sui risultati

**Conoscenze richieste**: Miglioramento continuo, relazioni con i Clienti, Contesto organizzativo

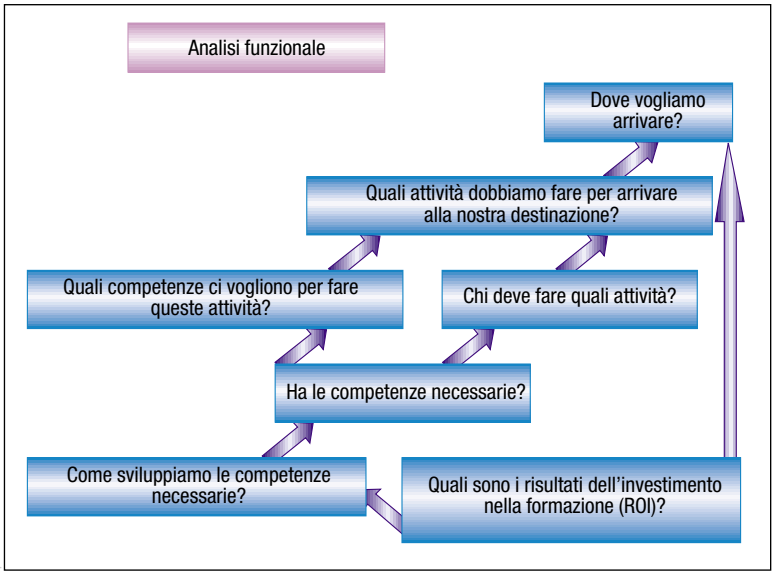
**Abilità richieste**: Affrontare i problemi e sfruttare le opportunità, Controllare la qualità del lavoro e confrontarla con i piani

Attraverso l'esame di quello che fanno i manager di successo, è stato possibile definire *criteri di prestazione* (o fattori critici di successo) per ogni attività. Questi *criteri di prestazione* permettono una valutazione ogget-

tiva dei risultati, e offrono uno strumento valido per la gestione della prestazione stessa (*performance*) e per l'allenamento/addestramento sul campo (*coaching*).
Per ogni attività sono anche state specificate le conoscenze, le capacità e le competenze che un lavoratore deve possedere, o sviluppare, se vuole svolgere il suo compito in modo competente. Queste specifiche permettono un' accurata valutazione delle competenze individuali e la preparazione di programmi di formazione e addestramento *on-the-job* mirati, efficaci ed economici.
In Inghilterra, i lavoratori vengono valutati da valutatori indipendenti, e, se dimostrano di essere competenti, ricevono un Certificato Nazionale Professionale (National Vocational Qualification), spendibile sul mercato del lavoro. Questi standard, quindi, facilitano la mobilità sia interna che esterna.
Una singola azienda o ente può fare la sua analisi funzionale e creare i suoi standard. Però gli standard nazionali sono tutti sviluppati da enti settoriali (per esempio: *Financial Services National Training Organisation, Central Government National Training Organisation*), vengono accettati da tutti le parti interessate, e sono accreditati dal *Department of Education and Skills*.
L'uso degli standard comporta cambiamenti radicali della cultura aziendale. Le agenzie del governo (per esempio l'*Inland Revenue*, l'ufficio delle entrate con 50.000 addetti) usano gli standard per creare una cultura di servizio al cittadino con un livello di servizio analogo a quello fornito dalle migliore aziende private. Le Casse di risparmio e le Assicurazioni mutue hanno trovato negli standard uno strumento che facilita la creazione di una cultura e linguaggio comune quando due o più casse si fondono. Le aziende della Grande Distribuzione Organizzata usano gli standard per ottenere una qualità di prestazione omogenea attraverso manager e commesse in migliaia di negozi geograficamente diffusi.
*Si puo' imparare qualcosa dall'esperienza inglese ed adattare l'approccio delle competenze, basato sul rigore dell'analisi funzionale, al contesto ed al carattere italiano*? La risposta è SI.
Credo che ci sia in Italia, come peraltro in tutta Europa, la forte esigenza di migliorare il profilo di competitività delle imprese e della P.A. centrale e locale, per agire efficacemente in un contesto di globalizzazione. Il modello delle competenze e degli standard professionali collega organicamente le esigenze di sviluppo delle conoscenze individuali e delle organizzazioni con gli strumenti per l'apprendimento.
C'è ora una terza risposta al quesito sul "come" si affronta lo *skill gap*:

Esiste una metodologia, efficace e verificata, con cui analizzare le attività di chi lavora, identificare i "gap" di competenza, assicurando la coerenza con obiettivi e processi dell'Organizzazione. Si chiama "standard professionali" ed è stata sviluppata ed applicata con successo in UK negli anni '90.

**FIGURA 7**
*(Skillmatrix, 2001)*




### 5.2.3. L'ANALISI FUNZIONALE

Una metodologia molto efficace per capire e collegare tra di loro i diversi "componenti" di una organizzazione è l'"analisi funzionale". Essa permette di collegare in un quadro comprensibile ed integrato sei componenti:
1. la missione aziendale;
2. le attività chiave che permettono di realizzare la missione;
3. i ruoli organizzativi;
4. le competenze necessarie sottese ai ruoli;
5. il "gap" di competenze per ogni persona a cui è affidato il ruolo;
6. i percorsi di sviluppo per ciascun individuo per superare i "gap".

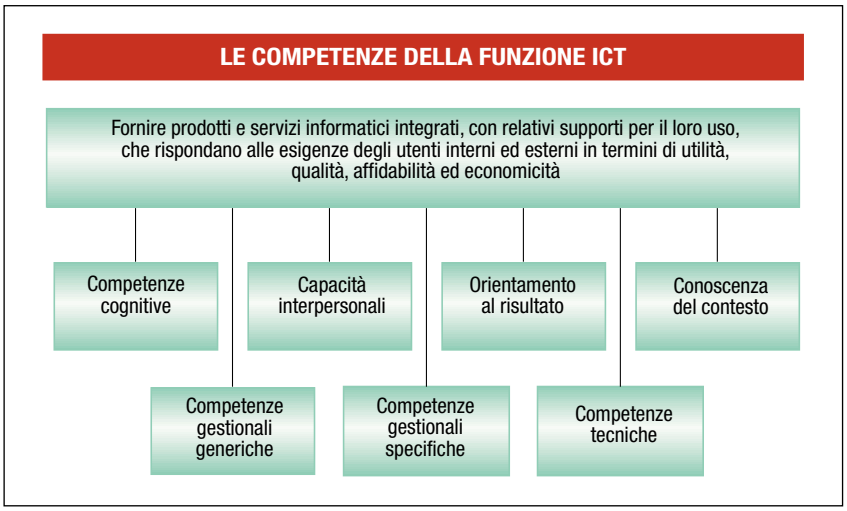
Nella figura 7 è esemplificata l'applicazione di questo metodo.

***Un esempio di applicazione del modello degli standard professionali e della "analisi funzionale" ad una intera funzione aziendale ICT.***

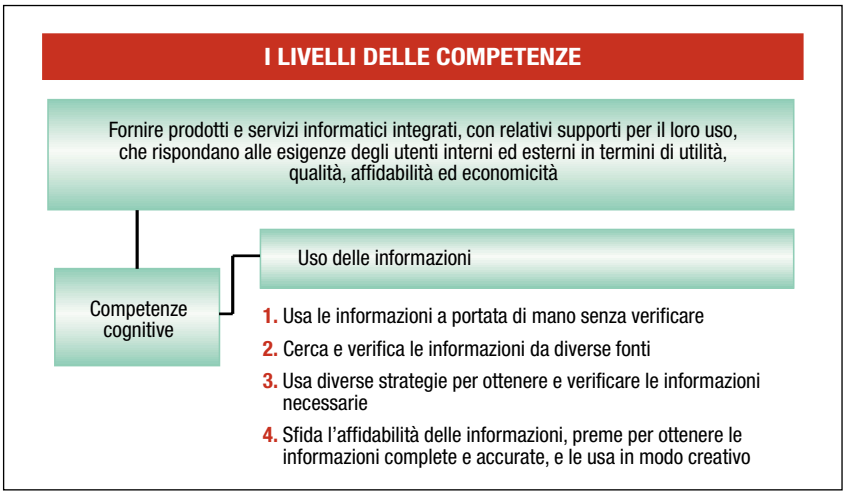
*Dove vogliamo arrivare?*


*Quali attività dobbiamo fare per arrivare alla nostra destinazione?*

Quali competenze ci vogliono per fare queste attività?




*A quale livello deve essere la specifica competenza?*
*1 basilare, 4 massimo*




## 5.3. La patente informatica ECDL

La Unione Europea, attraverso il CEPIS (*Council of European Professional Informatics Societies*), l'ente che riunisce le Associazioni europee di informatica, ha promosso uno strumento per l'alfabetizzazione di base degli utilizzatori delle tecnologie dell'informazione. Si chiama "***Patente europea di guida del computer***" (ECDL: *European Computer Driving Licence*). In Italia l'AICA [2] è l'ente nazionale di certificazione del programma ECDL.
È stato cioè definito uno standard per misurare le competenze di utilizzo dell'ICT che ha creato anche i percorsi formativi per raggiungere il livello base di conoscenze e di capacità. Ma che cosa significa realmente saper usare il computer? È sufficiente scorrere la lista degli esami da sostenere per ottenere la "patente" (vedasi scheda allegata) per rendersi conto che l'ECDL è una piattaforma di base di competenze, assolutamente necessaria per poter "navigare" nelle applicazioni informatiche e nella rete Internet, e poter usare efficacemente i più comuni strumenti di informatica individuale.
In base a un *protocollo di intesa con l'AICA*, il Ministero della Pubblica Istruzione ha adotta-

### COME SI OTTIENE LA PATENTE EUROPEA DEL COMPUTER?

Il candidato deve acquistare da un qualsiasi Centro accreditato (*Test Center o Test Point*) una tessera (*Skills Card*) su cui verranno via via registrati gli esami superati.
Gli esami sono in totale sette, di cui uno teorico mentre gli altri sono costituiti da test pratici. Il livello dei test è volutamente semplice, ma sufficiente per accertare se il candidato sa usare il computer nelle applicazioni standard di uso quotidiano. Più precisamente, sono previsti i seguenti moduli:
1 - Concetti teorici di base (*Basic concepts*)
2 - Uso del computer e gestione dei file (*Files management*)
3 - Elaborazione testi (*Word processing*)
4 - Foglio elettronico (*Spreadsheet*)
5 - Basi di dati (*Databases*)
6 - Strumenti di presentazione (*Presentation*)
7 - Reti informatiche (*Information networks*)
Ogni esame può essere sostenuto presso un qualsiasi Centro accreditato in Italia o all'estero. Il candidato non è, cioè, obbligato a sostenere tutti gli esami presso la stessa sede e inoltre può scaglionarli nel tempo (la tessera ha una validità di tre anni).
Quando ha superato tutti gli esami, egli riceve la patente (diploma) da parte dell'ente nazionale autorizzato ad emetterla (in Italia, l'AICA). Ogni esame può essere sostenuto presso un qualsiasi Centro accreditato in Italia o all'estero. Il candidato non è, cioè, obbligato a sostenere tutti gli esami presso la stessa sede e inoltre può scaglionarli nel tempo (la tessera ha una validità di tre anni).
A maggio 2001 AICA ha introdotto un nuovo sistema, ALICE, che automatizza in modo integrale gli esami che abilitano al rilascio del diploma ECDL. Per maggiori informazioni su ALICE si veda la pagina
http://www.aicanet.it/alice/alice.htm
informazioni generali: http://www.aicanet.it/ecdl.htm
materiale didattico: http://www.aicanet.it/materiale_didattico.htm

to ECDL come standard per la certificazione delle competenze informatiche nella scuola. Di conseguenza la patente europea del computer è accettata, senza problemi, come credito formativo negli esami di stato per il diploma di maturità.

## 6. UN CASO EMBLEMATICO: QUALE E QUANTA LOGISTICA PER L'E-BUSINESS

Un gruppo di lavoro dell'AILOG ([3] Associazione italiana di logistica e di supply chain management) ha analizzato nel 2001 l'interazione tra le tecnologie ICT, in particolare l'e-Business, e la Logistica giungendo a conclusioni che sono molto interessanti per il ragionamento che stiamo conducendo sullo "*skill shortage*", che diventa sempre più "*skill gap*". Sono stati estratti nove messaggi logistici, ossia nove sfide che le nuove tecnologie da un lato "creano" perché generano cambiamento e dall'altro "risolvono" offrendo soluzioni.
Il messaggio centrale è molto semplice e diretto: **e-Business è soprattutto Logistica e sempre più Logistica**. La Logistica *non è solo lo strumento che abilita il business sulla rete Internet*, è qualcosa di più, ossia sempre di più la Logistica è il modello di business, è la business idea. Si evidenzia la centralità della competenza Logistica ai diversi livelli di progettazione e gestione: nuovi modelli di impresa virtuali, nuovi modelli di business logistico-centrici, creazione del valore nella supply chain, gestione a distanza di processi di codesign e comakership.
Logistica e ICT insieme stanno innescando dei cicli di sviluppo importanti.
**Si devono creare nuove competenze per progettare e governare i sistemi di e-Business e in generale l'innovazione**.
Come possiamo leggere questi messaggi? Secondo me la lettura principale è che i confini delle competenze, nello specifico tra Logistica ed ICT, è sempre meno definibile; sempre di più emerge ***la competenza della "innovazione" che è trasversale***. Queste capacità di "trasversalità", cioè di lavorare insieme, di integrazione di conoscenze, diventano essenziali per trovare idee e soluzioni ai problemi; lo "skill gap" sta diventando quasi un fatto costitutivo della sfida della integrazione

delle nuove tecnologie nell'economia e nei modelli correnti di business.
La quarta risposta alla domanda iniziale dell'articolo potrebbe proprio essere questa:

> Se vogliamo veramente sfruttare il potere abilitante delle nuove tecnologie bisogna sviluppare nuove competenze, non solo tecniche ma anche manageriali, nuovi profili professionali, nuove modalità di apprendimento.

## 7. CONCLUSIONI

Mi piace citare Gian Filippo Cuneo [10] che scrive l'11 aprile 2001 nella sua mailing "Pensieri laterali":
*"Internet è qui per restare; chi prende sollievo da qualche fallimento di impresa virtuale vuole solo allontanare da se' la fastidiosa inevitabilità del cambiamento che tocca tutte le imprese. Accettare l'era di internet, evitare i pericoli di una risposta ritardata e svogliata, e coglierne le opportunità è prima di tutto una questione di schemi mentali di riferimento; le imprese italiane hanno normalmente schemi vecchi, inadatti all'era di Internet.....L'imprenditore che non ha ancora capito il fenomeno Internet pensa che l'on-line sia un fatto eccezionale, che l'impresa debba essere integrata, che sia importante produrre prodotti invece che servizi, che il servizio debba essere fatto da umani e non da computer, che sia sufficiente una segmentazione della clientela approssimativa, che controllare il capitale sia sufficiente a gestire imprese basate sulla conoscenza, che esista un mercato domestico, e che basti fare profitti per vincere la concorrenza e mantenere il valore dell'impresa. L'errore primigenio sta nel pensare che il mondo di ieri sia anche quello di domani; purtroppo, però, il futuro non e' più quello che era."*
***Lo skill gap piuttosto che lo skill shortage è la vera sfida nei prossimi anni.***
Lo *skill gap*, parafrasando scherzosamente la citazione di cui sopra, "*è qui per restare ed è prima di tutto una questione di schemi mentali di riferimento*".
Gli standard professionali, le metodologie per rilevare le attività e le competenze, e le stesse tecnologie ICT, ci possono peraltro aiutare moltissimo ad impostare un buon piano d'azione per affrontare questa "emergenza, che sta diventando continua".
La principale raccomandazione è quella di dotarsi di un buon schema di riferimento, cioè di un modello che renda chiari ed espliciti i collegamenti tra i bisogni ed i desideri degli individui e gli obiettivi e gli schemi di lavoro delle organizzazioni in cui questi individui operano. Il segreto potrebbe proprio essere quello di "vedere" ed "agire" in modo integrato: integrazione tra individui ed organizzazione, integrazione tra discipline.

## Bibliografia


[1] AAVV: *Harvard Business Review on Knowledge Management*. Harvard Business School Press, 1998.

[2] Aica: *La patente informatica ECDL*, http://www.aicanet.it/ecdlwhat.htm

[3] Ailog: *Gruppo di lavoro Logistica&e-Business, anno 2001*, www.ailog.it

[4] Anasin-Federcomin: *ict-job, Internet*, http://www.ict-job.it/

[5] Assinform: *Rapporto sull'informatica e le Telecomunicazioni, anni 1998, 1999, 2000, 2001*, www.assinform.it

[6] Assinform: *Rapporto 2001 sull'occupazione nel settore dell'Informatica e delle telecomunicazioni in Italia*, www.formazioneict.assinform.it

[7] Charles Leadbeater: *Living on thin air*. Penguin Books, 2000.

[8] Federcomin, Rapporto: *Occupazione e formazione nella Net economy*, 2000, 2001, www.federcomin.it

[9] FidaInform: rivista n. 22, settembre 2001, intervista a Giovanni Degli Antoni, Skill shortage, un falso problema, www.fidainform.org

[10] Gianfilippo Cuneo: *Pensieri laterali*,11 Aprile 2001, pensieri_laterali@it.buongiorno.com

[11] Trevor Boutall: *The good Manager's guide, MCI*, London, 1997.



RENZO PROVEDEL, ingegnere, ha sviluppato competenze di management in grandi aziende (Fiat, Iveco, ILVA) nei settori Informatica, Telecomunicazioni, Organizzazione e Logistica, ricoprendo responsabilità a livello Corporate, divisionale ed internazionale. Oggi è Imprenditore della net economy, titolare di un'azienda di servizi alle imprese per l'applicazione delle tecnologie ICT in processi di innovazione e di riposizionamento strategico. Presidente di FareImpresa, consigliere di AICA, di FidaInform, di AILOG.
E-mail: provedel@fareimpresa.net

# PARADIGMA ERP E TRASFORMAZIONE DELL'IMPRESA

**Gianmario Motta**

Il software ERP (Enterprise Resource Planning) hanno trasformato il sistema informativo aziendale. L'articolo illustra in primo luogo quelle caratteristiche distintive del paradigma ERP, che ne hanno favorito il successo: unicità del dato, modularità, prescrittività. Successivamente, discute le trasformazioni che gli ERP hanno indotto nel funzionamento delle imprese ai livelli dei processi operativi, manageriali, interaziendali. Infine considera i benefici potenziali offerti dalla trasformazione e propone uno schema di misurazione del suo valore, basato sulla valutazione dei vantaggi d'efficienza e d'efficacia.

## 1. LA SUITE ERP

L'acronimo ERP (*Enterprise Resource Planning*) è stato coniato agli inizi degli anni Novanta dal Gartner Group per indicare una *suite* di moduli applicativi integrati che supportano l'intera gamma dei processi di un'impresa. Oggi, una *suite* completa comprende decine di moduli applicativi, che possono essere schematicamente classificati nei tre gruppi di moduli *core* settoriali, moduli *core* intersettoriali e moduli *extended*.
Uno schema generale della *suite* ERP è in figura 1, dove la *suite* ERP è rappresentata da uno schema a T. I moduli settoriali sono la gamba della T, in quanto rappresentano la verticalizzazione delle applicazioni in ogni singolo settore industriale. La barra è formata dai moduli intersettoriali, in quanto orizzontali rispetto ai settori industriali, e dai moduli extended, che integrano l'azione degli ERP verso i mondi dei clienti e dei fornitori. La figura riporta anche un sommario elenco dei titoli dei principali moduli, che sono illustrati qui di seguito.

### 1.1. Moduli core intersettoriali

I moduli *core* intersettoriali sono sostanzialmente invarianti rispetto ai singoli settori industriali; in generale, informatizzano le attività aziendali di supporto.
I moduli istituzionali, così detti in quanto riflettono la regolamentazione pubblica, servono le attività amministrative, come la contabilità civilistica, la contabilità gestionale e la finanza aziendale, e la gestione delle risorse umane, che include sia le procedure contabili delle paghe sia i processi di gestione e sviluppo del personale.
I sistemi direzionali comprendono una serie di moduli che servono i processi, appunto, di conduzione manageriale dell'impresa, come la pianificazione strategica, la programmazione ed il controllo del budget, l'analisi dei costi e, più in generale, il reporting aziendale.
Sono intersettoriali, in quanto trasversali ai settori industriali, anche i moduli applicativi che pianificano e controllano le attività dei progetti ed elaborano la contabilità degli investimenti.
Fra i moduli intersettoriali includiamo anche il

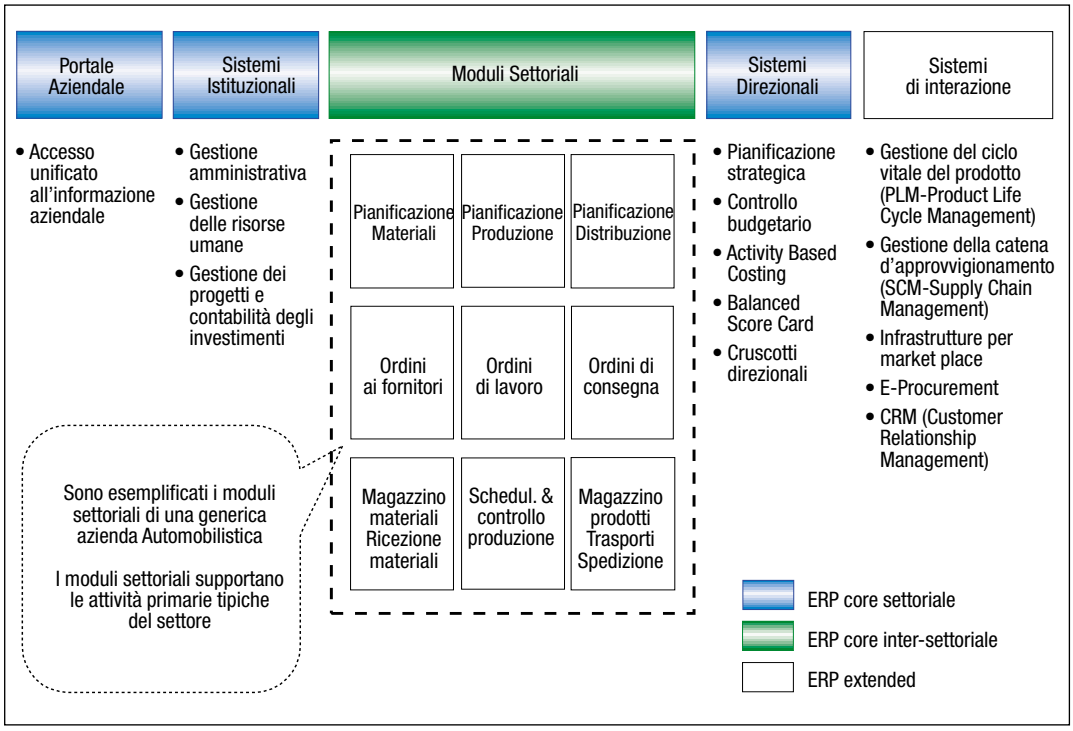



**FIGURA 1**
*La suite core ed extended*

portale aziendale. Apparso alla fine degli anni Novanta, offre un accesso via Web alle informazioni aziendali.
L'alta uniformità dei moduli intersettoriali ha favorito l'industrializzazione dell'offerta ERP in termini di qualità/costo; per contro, il cuore del sistema informativo aziendale, quindi il mercato maggiore, è dato dai moduli settoriali.

## 1.2. Moduli core settoriali

Le *suite* settoriali comprendono normalmente i moduli che supportano le attività primarie dell'azienda, tipiche del settore; sono molto diversificate, poichè riflettono le peculiarità d'ogni settore. Nella figura 1 è esemplificata, in sintesi estrema, la *suite* del settore automobilistico. Essa comprende una serie di moduli, che servono i processi d'approvvigionamento, produzione e vendita ai livelli di pianificazione, gestione degli ordini, attività fisiche.
La *suite* del settore automotoristico è diversa da quella del settore elettrico, che invece comprenderà moduli per la programmazione e controllo dei lavori (allacciamenti, nuovi impianti, dismissioni ecc.) e per la bollettazione.
Le *suite* settoriali sono numerose (per esempio il *vendor* SAP ne elenca una ventina) ed è conseguentemente elevato il costo sostenuto dai *vendor* per concepirle, realizzarle e mantenerle nel tempo. Non sorprendentemente, l'effettiva completezza delle *suite* settoriali è piuttosto variabile. In generale, è massima nel settore automotoristico, terra d'origine degli ERP, mentre è più limitata in altri settori, come banche o pubblica amministrazione, dove le *suite* includono ancora un limitato numero di moduli.

## 1.3. Extended ERP

L'*extended* ERP è formato da una serie di moduli che gestiscono le transazioni interaziendali e, più in generale, l'interazione fra più aziende o fra una singola azienda e clienti o fornitori.
In generale, queste *suite* supportano il ciclo vitale del prodotto (PLM, *Product Lifecycle Management*), la catena d'approvvigionamento (nota come SCM, *Supply Chain Management*), le interazioni con il cliente (CRM, *Customer Relationship Management*), l'*E-Procurement* e forniscono infrastrutture informatiche ai cosiddetti *Market Place*. Queste suite sono apparse sul mercato a partire dal 1995 come applicazioni indipendenti e separate dagli ERP core; con gli anni 2000, sono state integrate.
Valore distintivo dello schema ERP *extended* è l'integrazione fra transazioni interaziendali e transazioni interne. Per esempio, le *suite* ERP

| Numero di installazioni (I) o clienti (C) | Vendor | | |
|---|---|---|---|
| | Oracle (C) | People soft (C) | SAP (I) |
| **1999** | | | |
| WW | (ND) | 4.000 | 20.415 |
| Europa | (ND) | 900 | 13.500* |
| Italia | (ND) | (ND) | 517 |
| **2000** | | | |
| WW | 10.000 | 4.700 | 28.968 |
| Europa | 3.600 | 1.000* | 19.300* |
| Italia | 180 | 40 | 807 |
| **2001** | | | |
| WW | (ND) | 5.700* | 38.251 |
| Europa | (ND) | 1.300* | 25.500* |
| Italia | (ND) | 50 | 1.256 |

* Stime basate sulle dichiarazioni dei vendor

**TABELLA 1A**
*Diffusione dei maggiori vendor ERP*

| Vendor | Vendite (Milioni $) | Quota mercato |
|---|---|---|
| SAP | 5.839 | 30% |
| Oracle | 2.870 | 15% |
| Peoplesoft | 1.736 | 9% |
| J.D.Edwards | 980 | 5% |
| Altri | 7.127 | 41% |

**TABELLA 1B**
*Vendite 2000 dei maggiori vendor ERP (da AMR 2001)*

*core* sono integrate con i moduli CRM, che gestiscono i canali di contatto con il cliente (*call-center*, internet, agenti, negozi). L'integrazione fra CRM ed ERP *core* assicura l'effettiva esecuzione delle richieste del cliente (p.e. l'ordine di un'automobile, raccolto dai sistemi CRM, è programmato e controllato dai sistemi ERP *core* che gestiscono la produzione e la distribuzione). In altri termini, i sistemi CRM e, in generale, i sistemi d'interazione formano il *front-end* della azienda verso clienti e fornitori, mentre i sistemi ERP *core* formano il *back-end*.

## 1.4. Diffusione

Con l'estensione dello schema ERP, le aziende quindi hanno a disposizione una gamma molto ampia d'applicazioni informatiche; infatti:

- le *suite core* informatizzano le attività aziendali interne di livello operativo e direzionale;
- le *suite extended* informatizzano le transazioni interaziendali verso fornitori e clienti.

Non sorprendentemente, gli ERP sono divenuti uno dei massimi mercati d'applicazioni software. Alla fine degli anni Novanta, nel mondo si spendevano annualmente oltre 23 miliardi di dollari e i sistemi ERP erano installati in oltre 30.000 imprese. Oltre il 50% delle aziende europee ha uno o più moduli ERP e oltre il 35% li usa in tre o più aree funzionali [11]. Il numero totale di *vendor* sul mercato può essere stimato prossimo al centinaio. Tuttavia, solo alcuni, fra cui SAP, Oracle, Peoplesoft e JDEdwards, sono in grado di offrire l'intera gamma ERP. Il loro scarso numero riflette gli enormi investimenti necessari allo sviluppo ed al mantenimento di una gamma completa di moduli, intersettoriali e settoriali (Tabella 1).
Terza azienda software del mondo, SAP è nata intorno all'idea ERP e ha una quota stimata intorno ad un terzo del mercato. Il primo prodotto SAP, R/1, è stato un software *batch* per *mainframe*, sostituito nel 1981 da R/2 *on-line* e, nel 1992, da R/3 *on-line* e su architettura client-server, che negli anni 2000 è stata integrata da architetture *web-enabled*. A fine 2001, SAP, leader di mercato, dichiarava 12 milioni d'utenti, 44.500 installazioni, 1.000 aziende in partnership di vario tipo, 21 suite settoriali.
Oracle lancia la sua *suite* ERP negli anni Novanta, iniziando anch'essa dal core ERP poi integrato a portali e CRM, sfruttando ampiamente le proprie tecnologie di basi dati e In-

ternet. Punto di partenza di Peoplesoft sono i sistemi di gestione delle risorse umane, cui si aggiunge, dalla metà degli anni Novanta, la filiera standard ERP e CRM. JD Edwards, fondata nel 1977, si rivolge più specificatamente alla media impresa con un'offerta piuttosto ampia (6.000 clienti, di cui 300 in Italia).
Ai *vendor* multinazionali a gamma completa, si affiancano qualche decina di *vendor* multinazionali focalizzati su gruppi di moduli o su estensioni dello ERP *core*. Per esempio, Siebel, *leader* nel CRM, ha sviluppato un'offerta completa di CRM, che è a sua volta verticalizzata in soluzioni settoriali ed integrata con suite ERP. A sua volta, SAS, azienda *leader* di software statistici, offre una suite di CRM analitico, che elabora analisi sul comportamento del cliente o supporta la pianificazione delle campagne di promozione.
I grandi *vendor* multinazionali dominano nella grande impresa. Nelle piccole e medie imprese, invece, i primi 5 *vendor* hanno solo una quota minoritaria del mercato.

## 2. IL PARADIGMA ERP

La *suite* ERP rispecchia una precisa concezione del sistema informativo aziendale, con caratteristiche distintive:
- unicità dell'informazione;
- estensione e modularità funzionale;
- prescrittività.

Queste caratteristiche formano un paradigma funzionale, che indichiamo come paradigma ERP (un analogo concetto di paradigma ERP è discusso in [11]).

### 2.1. Unicità dell'informazione

Gli ERP sono caratterizzati da una base dati unica. Unica fisicamente od unificata attraverso un comune *repository* dei dati e servizi di replica automatica, memorizza i dati condivisi intorno alla quale ruotano i moduli. La base dati unica è una conquista sostanziale degli ERP che ha molti ed importanti vantaggi (Figura 2).
In primo luogo, l'aggiornamento unificato delle basi dati abilita la sincronizzazione di processi gestionali interdipendenti: p.e. l'arrivo di un materiale al magazzino aggiorna la situazione delle scorte, degli ordini ai fornitori e della contabilità dei fornitori, dando ai corrispondenti processi un'informazione unica e sincronizzata. Ciò non è possibile nelle tradizionali architetture ad isole, dove le basi dati sono separate e i dati comuni sono sincronizzati attraverso periodici processi d'allineamento. Nell'architettura ad isole, le informazioni sullo stesso oggetto (cliente, fornitore o materiale) sono temporalmente sfasate e ridondanti: il mancato pagamento di un cliente può non essere notificato in tempo alla gestione degli ordini, così che la situazione del cliente alla gestione degli ordini contrasta con quella della contabilità clienti (Figura 3).

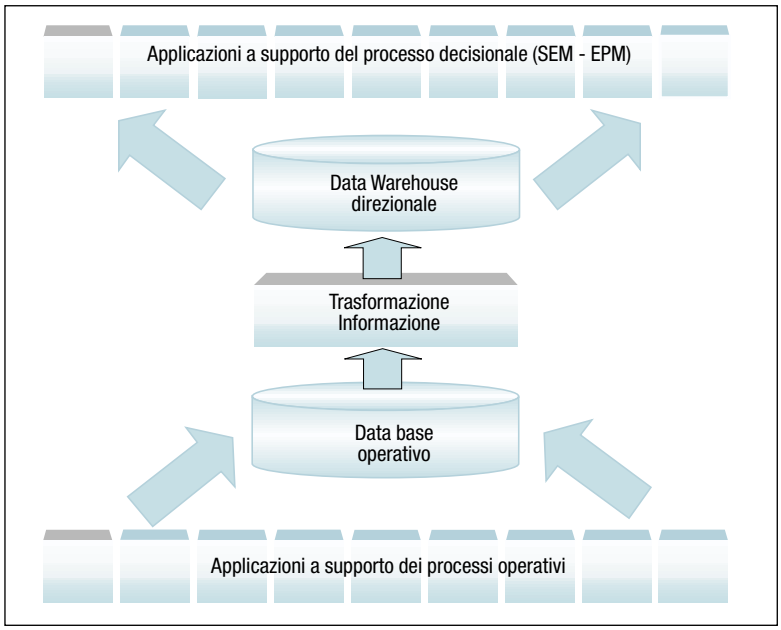



**FIGURA 2**
*Architettura ERP dei dati*

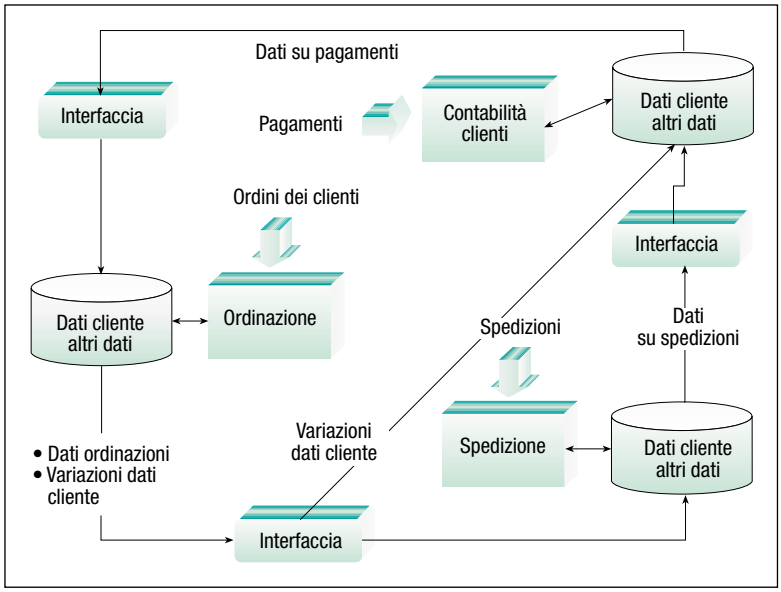



**FIGURA 3**
*Architettura ad isole di sistemi legacy: le basi dati sono sincronizzate da interfacce dedicate che propagano le modifiche*

In secondo luogo, l'architettura ERP certifica l'informazione e ne garantisce la tracciabilità: ogni evento di un processo per esempio la gestione di un magazzino, è testimoniato da un documento, per esempio una bolla di prelievo, che è specificatamente registrato nella base dati; ogni evento gestionale si riflette in una variazione di stato della base dati e la variazione é certificata da un documento.
Infine, l'unicità della base dati a livello operativo favorisce, in modo del tutto naturale, l'unicità dei dati per la direzione aziendale. L'unicità è ottenuta attraverso l'integrazione verticale dell'informazione operativa e dell'informazione manageriale. L'integrazione verticale si basa su un *Data Warehouse*, che memorizza i dati, aggregati e trasformati, estratti dalla base dati operativa (e da altre fonti). I dati sono elaborati da una *suite* di applicazioni, dette SEM (*Strategic Enterprise Management*) od EPM (*Enterprise Performance Management*), che assistono il management nella formulazione della strategia, nel *budgeting*, nell'analisi dei risultati.
L'integrazione rende disponibili informazioni sintetiche univoche (in quanto basate su dati operativi univoci ed unici), con un vantaggio rilevante per il management. Come molti studi hanno notato, la qualità dei dati è fra i vantaggi più apprezzati delle soluzioni ERP.

## 2.2. Estensione e modularità

Grazie all'estensione molto ampia, la *suite* ERP si propone come soluzione di riferimento per il sistema informativo aziendale, nelle sue componenti intra-aziendale, operativa direzionale, ed inter-aziendale.
Tuttavia, l'estensione funzionale sarebbe vana se la *suite* non fosse composta da moduli autosufficienti. Grazie alla modularità, l'azienda può scegliere una strategia d'implementazione coerente con la situazione dei sistemi e con il grado di rischio che è in grado di sostenere.
Una diffusa strategia semplice ed a basso rischio è l'implementazione parziale: l'azienda, cioè, sceglie di realizzare un piccolo numero di moduli, che vanno a sostituire preesistenti sistemi *legacy*.
La strategia, più ambiziosa, di implementare un elevato numero di moduli può essere attuata in due varianti, *one stop shopping* e *best of the breed*. Nel primo caso, privilegiando linearità e semplicità, l'azienda usa i moduli di un solo *vendor*, mentre, nel secondo, mette insieme moduli di più *vendor*, alla ricerca della soluzione ottimale per ogni processo aziendale, p.e. scegliendo il *vendor* A per la gestione del personale ed il *vendor* B per la gestione amministrativa.
Osserviamo che, data la modularità e l'ampia estensione funzionale degli ERP, la progettazione diventa simile ad una specie di *LEGO*, in cui è critico l'incastro fra i diversi moduli ERP, magari di più fornitori, e fra i moduli ERP e gli eventuali moduli *legacy*. Infatti, vanno garantite l'unicità e la sincronizzazione delle informazioni, attraverso interfacce standard, API (*Application Programming Interface*) e software di workflow o d'integrazione.

## 2.3. Prescrittività

I moduli ERP incorporano, in misura significativa, una logica di processo gestionale. Per esempio, la transazione di ricevimento dei materiali a magazzino presuppone un ordine al fornitore: un materiale non entra in azienda se non è stato ordinato e non può essere ordinato se non è stato richiesto da un ente aziendale autorizzato. Il software quindi norma il comportamento dell'utente aziendale, ribaltando la tradizionale concezione secondo cui è il software che si deve adattare all'utente.
La prescrittività ha importanti implicazioni. In primo luogo, semplifica l'analisi dei requisiti. L'analista, infatti, non deve specificare tutto il processo gestionale e tutto il sistema, ma si concentra sulle differenze rispetto al modello standard: definisce il processo ottimale, lo incrocia con le funzioni del sistema e sceglie le funzionalità del sistema. La progettazione funzionale diventa quindi una sorta di attività "taglia ed incolla" su di un *menu* di opzioni predefinite.
In secondo luogo, la prescrittività favorisce la standardizzazione dei processi ed uniforma i comportamenti, un vantaggio rilevante per le aziende distribuite territorialmente e le multinazionali.
Infine, la prescrittività può favorire una razionalizzazione dei processi, facendo coincidere il progetto informatico ERP con un progetto di

razionalizzaione organizzativa, meglio noto sotto la sigla BPR (*Business Process Reengineering*).
Tuttavia, la prescrittività comporta anche una certa rigidità, che può rendere gli ERP incompatibili con le specificità dell'utente. Infatti, se la razionalizzazione organizzativa richiesta è vasta e profonda, il progetto implica un costoso e rischioso intervento sul tessuto organizzativo dell'impresa. L'intervento può risultare infattibile per i tempi, troppo stretti, per i contenuti dei processi, incompatibili con il sistema dei valori esistente nell'azienda, per il rischio della trasformazione, troppo ampia, o, infine, per la mancanza di un gruppo di lavoro di quantità e qualità adeguata o, equivalentemente, per limiti di budget. L'alternativa ad adattare l'azienda al sistema è quella, piuttosto costosa e di un certo rischio tecnico, di adattare il sistema all'azienda, riscrivendo e/o modificando i moduli software.

## 3. TRASFORMAZIONE DELLA IMPRESA

### 3.1. La griglia di trasformazione

La diffusione degli ERP è evidente dai numeri ed è evidente la trasformazione delle imprese. Una volta lente, con una produzione inflessibile ed approvvigionamenti rigidi, in due decenni sono diventate capaci di gestire ordini personalizzati, lungo tutto il ciclo dal cliente finale al fornitore. E questo con scorte molto minori e una produttività molto maggiore. In generale, le caratteristiche degli ERP hanno contribuito a una serie di trasformazioni e queste trasformazioni, a loro volta, hanno generato, in varia misura, alcuni vantaggi. Le trasformazioni rilevabili appaiono riguardare i processi gestionali a diversi livelli (Figura 4):

- processi aziendali di livello operativo;
- processi aziendali di livello manageriale;
- processi interaziendali;
- modello di business (la trasformazione però appare molto più sfumata e controversa).

Le prime ricerche sulle trasformazioni aziendali indotte dalla IT sono degli anni Ottanta, con il riconoscimento del duplice ruolo della IT come tecnologia di produzione e come tecnologia di coordinamento [16]. La capacità delle IT, degli ERP in particolare, di trasformare i processi gestionali è discussa nella vastissima letteratura BPR (*Business Process Reengineering*), che segue lo storico articolo di Hammer "*Don't automate, obliterate*" [9] e si può dire conclusa da un articolo di Davenport [5] dal significativo titolo "*Putting the enterprise into the enterprise system*".

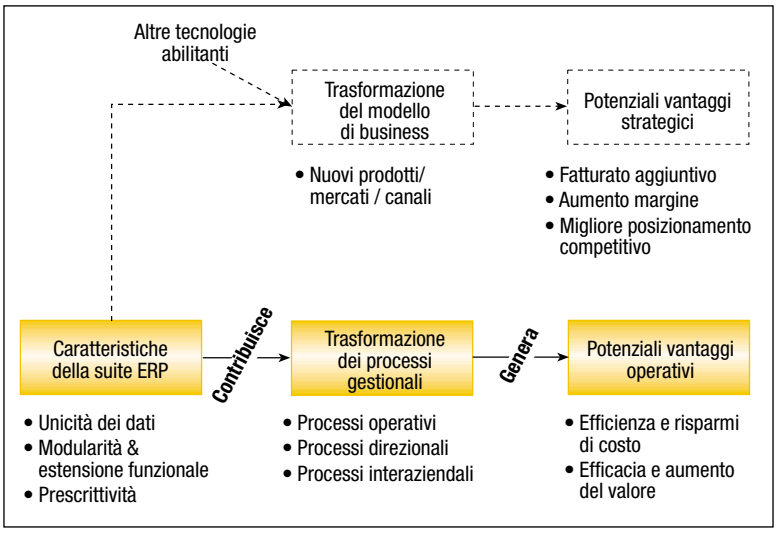



**FIGURA 4**
*La catena di trasformazione dei processi gestionali*

Vari autori hanno proposto schemi di correlazione fra caratteristiche di applicazioni IT e/o ERP e trasformazione dei processi. Venkatraman [21] ha ipotizzato una serie di stadi di *business transformation*, che vanno dalla ottimizzazione localizzata delle attività esecutive, all'integrazione fra attività interne, alla riconcezione dei processi, per arrivare alla riconcezione delle relazioni esterne e dello stesso business aziendale. Ciascuno stadio comporta una trasformazione più ampia e profonda dello stadio precedente e maggiori benefici potenziali.
Qui di seguito consideriamo le caratteristiche delle trasformazioni dei processi e dalla trasformazione del modello di business.

### 3.2. Trasformazione dei processi operativi

Con "trasformazione dei processi operativi" intendiamo un cambiamento dei processi che ne migliora l'efficienza e l'efficacia.
Grazie alla loro prescrittività, gli ERP dovrebbero cambiare l'organizzazione aziendale e portare ad un'organizzazione "processiva", più agile e flessibile, non parcellizzata, orientata a dare valore al cliente, con personale

| Variabile | Evoluzione |
|---|---|
| Flusso delle attività | • Minor numero di passi<br>• Minore durata del processo |
| Organizzazione operativa | • Arricchimento delle mansioni<br>• "Deparcelizzazione" del flusso del lavoro |
| Personale | • Versatilità operativa |
| Sistema di incentivazione e di controllo | • Obiettivi di servizio al cliente<br>• Obiettivi di performance sui processi gestionali |

**TABELLA 2**
*Evoluzione processiva delle variabili organizzative (adattato da Bracchi & Motta 2001)*

versatile e polivalente, in grado di svolgere un ampio ventaglio d'attività (Tabella 2).
I risultati dei progetti ERP indicano che l'evoluzione processiva esiste, ma parziale e condizionata da una serie di fattori; più precisamente:
❏ la trasformazione dei processi avviene e ne riguarda sia l'efficienza sia l'efficacia [6 - 13 - 18];
❏ un approccio sistemico al cambiamento, insieme con un iter di progettazione cauto, graduale, burocratico possono portare al successo, in quanto minimizzano il rischio [14];
❏ per ottenere veramente la trasformazione sono necessari un orientamento informatico appropriato del management e un ben orchestrato lavoro d'attori organizzativi che facilitino il cambiamento [2]; L'organizzazione e la gestione, nel senso più ampio, del progetto sono cruciali per il successo o il fallimento dei progetti ERP [14]: il presidio deve essere esteso a tutte le fasi dei progetti ERP, dal lancio all'accettazione ed alla stabilizzazione [18];
❏ un'elevata trasformazione dei processi, connessa all'adozione di ERP, aumenta il rischio del progetto, e richiede il presidio di un'ampia gamma di fattori di successo [19].
In sintesi, la trasformazione, per avvenire, richiede, in aggiunta al progetto informatico, uno specifico progetto organizzativo. Va osservato che l'attuazione di una trasformazione organizzativa significativa richiede l'impegno del management, non scontato e comunque costoso. Conseguentemente, è conveniente affrontare i progetti ERP, ad alto rischio organizzativo ed a medio rischio tecnico, riducendo i rischi, graduando i tempi, selezionando l'innovazione e coordinando, attraverso un'opportuna metodologia integrata, le filiere d'attività del progetto: implementativa, organizzativa, e infrastrutturale [3].

### 3.3. Trasformazione dei processi direzionali

Il contributo degli ERP alla trasformazione dei processi direzionali sta appunto nel loro contributo a rendere più efficiente e/o efficace l'informazione in input al processo decisionale e/o il processo decisionale stesso.
Un concetto rilevante per posizionare il contributo degli ERP è quello di IPC (*Information Processing Capacity*), che esprime l'adeguatezza di una organizzazione ad elaborare le informazioni richieste dai propri obiettivi e dal contesto in cui opera [3]. La capacità di un'azienda di operare in situazioni d'incertezza ambientale e di gestire strutture complesse è proporzionale alla sua IPC. In alternativa, l'azienda può investire in "risorse cuscinetto" (*slack resources*) come le scorte, che assorbono l'incertezza e diminuiscono il fabbisogno informativo, ma peggiorano le prestazioni d'efficienza e d'efficacia [8].
L'adozione degli ERP, specificatamente della suite "Sistemi Direzionali", aumenta la IPC attraverso l'aumento di valore della informazione, sotto vari aspetti:
❏ *ampiezza del dominio informativo*: l'ERP abbraccia tutta la catena del valore aziendale ed è integrabile con fonti esterne, pubbliche (Internet) e di fornitori;
❏ *disponibilità ed accessibilità*: l'informazione, memorizzata in basi dati unificate, è distribuibile in modo semplice, attraverso Internet ed accessi wireless;
❏ *velocità*: gli ERP accelerano il processo di creazione della informazione direzionale;
❏ *utilizzabilità*: gli ERP producono informazione fruibile anche per processi direzionali di livello strategico, come la Balanced Score Card [10].
La migliore qualità e, in definitiva, il minore costo dell'informazione offre un'opportunità anche per la trasformazione del processo decisionale. Secondo alcune teorie [12],

l'IT può ampliare il ruolo del management operativo (*empowerment*) ed accorciare le linee gerarchiche, introducendo un ciclo di controllo più breve e basato sulla collaborazione (Figura 5).

### 3.4. Trasformazione dei processi interaziendali

I processi interaziendali includono un arco ampio di relazioni [21] che spazia dalle transazioni interaziendali, alla gestione interaziendale di processi (scambi d'informazioni), a processi interaziendali (p.e. sviluppo congiunto di un nuovo prodotto) a sistemi di conoscenza condivisa.
Tale ambito di trasformazione riguarda primariamente i moduli *extended* ERP.
In primo luogo i moduli che gestiscono le transazioni con fornitori e partner permettono di realizzare sistemi inter-organizzativi per la cooperazione fra più aziende mediante condivisione d'informazioni (*e-procurement*), governo di flussi interaziendali di fornitura (EDI, sistemi di tracciamento d'ordini e di spedizioni), reti per svolgere un compito o servizio comune (pool di manutenzione impianti). Anche in questo caso, per ottenere una reale trasformazione, il progetto implementativo va integrato con un progetto logistico-organizzativo e con una valutazione della fattibilità del *business model* correlato [22].
In secondo luogo, i moduli CRM modificano sia il modo con cui l'azienda risponde al cliente sia il modo con cui il cliente interagisce con l'azienda. La trasformazione avviene attraverso l'attivazione di nuovi canali e delle corrispondenti tecnologie, come telefono (*call –center*), E-mail, Web fisso e mobile e, infine, sistemi per l'automazione delle forze di vendita, che, pur pre-esistenti, sono stati rinnovati tecnologicamente ed integrati con gli altri moduli CRM.
Un'analisi della trasformazione CRM esce, per la sua vastità, dagli scopi del nostro articolo. Ci limitiamo a sottolineare, che, poichè la risposta alle transazioni con i clienti e i fornitori è data dai processi interni, l'integrazione fra moduli *ERP core* e moduli di interazione è un requisito d'efficacia altrettanto importante della funzionalità dei moduli CRM, SCM che informatizzano il *front-end* rispettivamente verso il cliente od il fornitore.

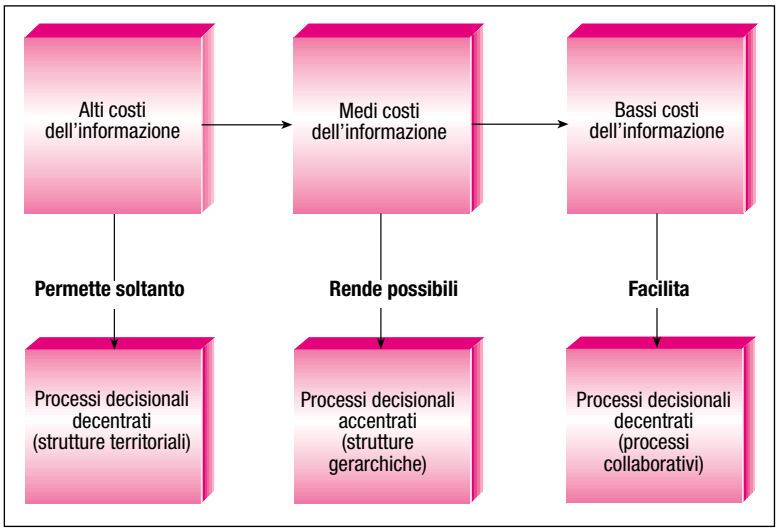



**FIGURA 5**
*Evoluzione dei processi decisionali nella teoria dell'empowerment di Malone*

### 3.5. Trasformazione del modello di business

Una trasformazione del modello di business implica la sostanziale innovazione del modello esistente. In sintesi, l'IT:
- rende possibili nuove attività con conseguenti nuovi ricavi;
- determina una domanda di nuovi prodotti e servizi;
- permette di sviluppare nuovi business a partire dai business esistenti.

Esempio significativo è Amazon, che innova la vendita per corrispondenza di libri e simili attraverso la tecnologia CRM. Un secondo esempio è la banca multicanale, che aggiunge alla tradizionale filiale i nuovi canali dei promotori finanziari, della banca telefonica (*call center*), della banca *on line* su *web*, sfruttando la capacità dei sistemi CRM di gestire in modo integrato il rapporto con il cliente. Un terzo esempio, emergente ora, è l'info-trattenimento: grazie alle nuove tecnologie di comunicazione (Fibre, ADSL) le aziende telefoniche offrono nuovi prodotti, come accesso Internet, intrattenimento e informazione. Con lo standard I-Mode, NTT introduce nella telefonia mobile il business dell'info-trattenimento, che raggiunge, a fine 2001, 30 milioni d'utenti.
Com'è evidente dagli esempi citati, la trasformazione è un'opportunità non di tutte le aziende ma solo di alcune, che sono connota-

te dall'elevata intensità informativa del prodotto e dei processi [16]. In secondo luogo, l'esistenza dell'opportunità non implica senz'altro il suo perseguimento: l'opportunità dell'e-business c'è per tutte le librerie, ma è stata perseguita soltanto da Amazon ed alcuni altri. Il perseguire un'opportunità riflette la strategia competitiva della singola azienda [2].
La trasformazione del modello di business passa attraverso i sistemi d'interazione con il cliente o con i fornitori. In questo caso, il contributo specifico degli *ERP core* è limitato. Tuttavia, un nuovo modello di business richiede anche (la progettazione di) un'efficiente catena del valore, che garantisca all'azienda la necessaria redditività. Tale requisito implica, a sua volta, una rivisitazione dei processi operativi interni ed una specifica strategia *ERP extended*. Amazon, che soltanto nel 2002 ha raggiunto il profitto, esemplifica quanto poco scontato sia il progetto operativo interno.

## 4. VANTAGGI E BENEFICI

La percezione comune associa ai progetti IT vantaggi d'efficienza e d'efficacia. Tuttavia, l'analisi aggregata dei dati di produttività d'aggregati nazionali od aziendali da' indicazioni contrastanti. Negli anni Ottanta, [20] evidenzia la non correlazione fra livello d'investimento IT e *performance* aziendale. Negli anni Novanta emerge la questione del *productivity paradox*, [4] che può essere così riassunto: gli investimenti IT salgono vertiginosamente mentre la produttività degli addetti (definita come quantità d'output prodotta per unità d'input) stagna o cresce marginalmente, come testimoniano le serie storiche USA e d'altri Paesi OCSE [7]. Il paradosso, tuttora irrisolto, é parzialmente spiegato dalla difficoltà di misurare in modo completo gli input e gli output: per esempio, l'investimento in Bancomat abbasserebbe la produttività delle banche se il loro costo fosse conteggiato nell'input ma il loro prodotto non fosse conteggiato nell'output (per esempio se la misura d'output fossero gli assegni per addetto).
Da parte nostra, ci sembra più coerente orientare la valutazione dei vantaggi sui processi, che sono lo scopo fondamentale degli ERP. Come già accennato, sono stati rilevati benefici specifici d'efficienza e d'efficacia, ottenuti con l'implementazione di sistemi ERP [6 - 13]. Tali benefici esprimono vantaggi operativi rilevanti su processi sia interni sia interaziendali (Tabella 3).
Una valutazione completa degli ERP deve quindi rispecchiare l'impatto complessivo sul business [17] ed includere anche significativi benefici intangibili, come l'accelerazione e la disponibilità dell'informazione manageriale [15].
Per quanto detto, riteniamo coerente para-

**TABELLA 3**
*Benefici ERP (rielaborato da [13])*

| Bene-fisico | ERP – core | ERP extended & SCM (Supply Chain Management) |
|---|---|---|
| Efficienza operativa | • Minori costi di staff (Business Process Reengineering - BPR)<br>• Minori scorte<br>• Minori costi logistici<br>• Minori costi di approvvigionamento<br>• Maggiore produttività e flessibilità | • Minori costi di staff (Business Process Reengineering - BPR)<br>• Minori scorte<br>• Minori costi logistici<br>• Minori costi di approvvigionamento<br>• Previsioni della domanda più affidabile |
| Efficacia operativa | • Maggiore tasso d'evasione ordini<br>• Migliorata capacità di risposta al cliente (*client responsiveness*) | • Maggiore tasso d'evasione ordini<br>• Migliorata capacità di risposta ai partner della catena di fornitura<br>• Minore *time-to-market* |
| IT | • Standardizzazione delle piattaforme IT | |
| Valore della Informazione | • Condivisione globale dell'informazione | |
| Altre | | • Creazione di nuove opportunità di mercato |

metrare la valutazione dei benefici sul valore della trasformazione dei processi. Il valore della trasformazione include il valore dei guadagni d'efficienza, ottenuti attraverso la migliore ingegneria dei processi, e dei guadagni d'efficacia, percepiti dai clienti, che comprendono il maggiore valore degli output (adottiamo le classiche definizioni Efficienza = Output effettivo / Input ed Efficacia = Output effettivo / Output atteso). In sintesi, il valore della trasformazione **T** può essere quantificato come la somma dei guadagni d'efficienza **E** e del guadagno d'efficacia **V** che è percepito dal cliente, cioè **T** = **E** + **V**. I benefici d'efficacia **V**, riflettono il guadagno di valore derivante da migliori prestazioni del processo, quali migliore livello di servizio e migliore qualità in uscita, e misurano la differenza fra il prezzo **P2** che il cliente è disposto pagare rispetto al prezzo **P1** ante-progetto, ovvero **V** = **P2** – **P1**. Si osservi che questo schema di valutazione della trasformazione è coerente con la lista di benefici citata in tabella 3; fanno eccezione le nuove opportunità di business, che richiedono un approccio specifico alla valutazione, a valle di un *business plan*.

## 5. PROSPETTIVE FUTURE

Dall'analisi dei *report* e delle presentazioni dei maggiori *vendor* ERP emergono una serie di linee evolutive generali [1 - 11].
Una prima ovvia linea d'evoluzione è tecnologica. Gli ERP, nati con architettura *client-server*, devono accedere ed essere accessibili via Internet e devono fornire portali di impresa. Questa evoluzione è inclusa nel paradigma standard degli ERP tracciato in figura 1.
Un'altra linea d'evoluzione è il completamento delle *suite* intersettoriali e settoriali, che permette ai *vendor* di allargare il mercato. Nell'ambito delle suite intersettoriali, lo sviluppo dei sistemi CRM, SCM e simili è ancora all'inizio; inoltre, sono da sviluppare i paradigmi gestionali di reale cooperazione interaziendale, non solo di scambio di dati. Nell'ambito delle soluzioni settoriali, molti settori sono ancora lontani da una *suite* completa, che comprenda tutti i sistemi d'elaborazione delle transazioni e raggiunga una completezza paragonabile alla *suite* per le aziende *automotive*.
Una terza linea d'evoluzione è l'integrazione. Pochi vendor riusciranno a sostenere il peso di una gamma completa ed integrata. Inoltre, molte aziende scelgono o sono costrette ad adottare soluzioni miste, in cui é necessario collegare *suite* ERP di fornitori diversi e sistemi ERP con sistemi *legacy*. Queste condizioni aprono un mercato ampio per i software d'integrazione delle applicazioni.
Un'ulteriore evoluzione è l'allargamento del mercato alle piccole e medie imprese attraverso *suite* semplificate e/o attraverso nuove modalità di fruizione. A questo scopo, alla modalità a progetto si affianca la modalità ASP (*Application Service Provider*), cui l'azienda accede via Web. L'azienda si limita a pagare l'uso di un software preconfezionato e verticalizzato, che risiede sui server di un centro servizi; in questo modo, può abbattere drammaticamente i costi d'utilizzo ed evitare i costi, i tempi ed i rischi dei progetti.

## 5. CONCLUSIONI

Il fenomeno ERP rispecchia la progressiva uniformazione del sistema informativo aziendale ed interaziendale in un paradigma completo ed integrato che ha trasformato, in varia misura, le aziende che lo hanno adottato.
La prima sostanziale trasformazione è in realtà la trasformazione del sistema informativo aziendale, che da collezione d'applicazioni diverse ed indipendenti diviene un'ordinata catena di montaggio e distribuzione dell'informazione.
La trasformazione del sistema informativo può favorire la trasformazione dei processi a livello operativo, direzionale, interaziendale. La trasformazione direzionale é un'area di potenziale alto successo, che integra bene modelli manageriali maturi con una tecnologia adeguata. La trasformazione interaziendale è ancora agli inizi. Il contributo degli ERP all'innovazione del modello di business appare limitato e incidentale.
Le trasformazioni qui discusse sono potenziali. La trasformazione effettiva è funzione dell'effettiva capacità dell'azienda di sfruttare le potenzialità ERP attraverso un'opportu-

na trasformazione del tessuto organizzativo ed un approccio cauto e ben bilanciato al progetto ERP.
Per misurare i benefici, potenziali ed effettivi degli ERP, è conveniente considerare il valore totale della trasformazione, concepito come la somma algebrica dei vari guadagni d'efficienza operativa e d'efficacia indotti dai progetti ERP.

## Bibliografia


[1] Austin RD, Cotteler MJ, Escalle CX: *Enterprise Resource Planning: Technology Note*. Harvard Business School, Publ. 1999. update nov 2001.

[2] Benjamin RI, Markus ML: The magic bullet theory in IT-enabled transformation. *Sloan Management Review*, 1997.

[3] Bracchi G, Francalanci C, Motta G: *Sistemi informativi e aziende in rete*. McGraw Hill, Milano, 2001.

[4] Brynjolfsson E, Hitt LM: Beyond the Productivity Paradox. *Communications of the ACM*, Vol. 41, n. 8, 1998, p. 49-55.

[5] Davenport TH: Putting the enterprise into the enterprise system. *Harvard Business Review*, Vol. 75, n. 4, 1998.

[6] Deloitte Consulting: ERP's second wave, European Research Presentation. *Deloitte Consulting*, 1999.

[7] Dewan S, Kraemer KL: International dimensions of productivity paradox. *Communications of the ACM*, Vol. 41, n. 8, 1998, p. 56-62.

[8] Galbraith JR: *Designing Complex Organizations*. Addison-Wesley Publishing Company Inc., 1973.

[9] Hammer M: Reengineering Work: Don't Automate, Obliterate. *Harvard Business Review*, Vol. 68, n. 4, 1990, p. 104-112.

[10] Kaplan RS, Norton DP: Using the Balanced Scorecard as a Strategic Management System. H*arvard Business Review*, Vol. 74, n. 1, 1996, p. 75-86.

[11] Mabert Vincent A, Soni Ashok, Venkatraman MA: Enterprise Resource Planning: common myths versus evolving reality. *Business Horizons*, May-June 2001, p. 69-75.

[12] Malone TW: Is empowerment just a fad? Control, management, and IT. *Sloan Management Review*, Vol. 38, n. 2, 1997.

[13] Ming-Ling C, Shaw WH: *Distinguishing the critical success factors between e-commerce, enterprise resource planning, and supply chain management*. Engineering Management Society, 2000, p. 596-601.

[14] Motwani J, Mircahandani D, Madan M, Gunasekaran A: Successful implementation of ERP projects: evidence from two case studies. *International Journal of Production Economics*, Vol. 75, 2002, p. 83-96.

[15] Murphy KE, Simon SJ: *Using cost benefit analysis for enterprise resource planning project evaluation: a case for including intangibles*. Proceedings of the 34th Annual Hawaii International Conference on System Sciences, 2001, p. 2955 -2965.

[16] Porter ME, Millar VE: How Information Gives You Competitive Advantage. *Harvard Business Review*, Vol. 63, n. 4, 1985, p. 149-161.

[17] Reneyi D, Sherwood-Smith M: Outcomes and benefit modeling from information systems investment. *The International Journal of Flexible Manufacturing Systems*, Vol. 13, 2001, p. 105-129.

[18] Ross JW, Vitale MR: The ERP revolution: surviving versus thriving. *Information Systems Frontiers*, Vol. 2, n. 2, 2000, p. 233- 241.

[19] Somers TM, Nelson K: *The impact of critical success factors across the stages of enterprise resource planning implementations*. 34th Hawaii International Conference on Systems Sciences, 2001.

[20] Strassmann PA: *The information pay-off*. The Free Press, 1985.

[21] Venkatraman N: IT enabled business transformation: from automation to business scope redefinition. *Sloan Management Review*, Vol. 35, n. 2, Winter 1994, p. 74-88.

[22] Volkoff O, Chan YE, Newson PEF: Leading the development and implementation of collaborative interorganizational systems. *Information and Management*, Vol. 35, 1999, p. 63-75.



GIANMARIO MOTTA Laureato in filosofia, è stato dirigente e partner in multinazionali del settore ITC, fra cui EDS e Deloitte Consulting. Studioso di sistemi informativi direzionali e delle strategie aziendali IT, è docente di sistemi informativi al Politecnico di Milano, ed è autore, con Giampio Bracchi, di numerosi testi.
E-mail: gianmario.motta@polimi.it

# CARENZA DI COMPETENZE: "SKILL SHORTAGE" NEL SETTORE ICT

**Renzo Provedel**

L'articolo propone una lettura ampia e critica del fenomeno della carenza di competenze ICT e risponde a due domande: esiste ancora, in questo quadro economico mondiale incerto, un fabbisogno così rilevante di competenze e di specialisti ICT, e se sì quali sono le possibili azioni concrete per affrontarlo? Sono valutati l'urgenza dell'intervento, la metodologia ottimale per far crescere le competenze ICT nelle organizzazioni e se le nuove professioni ICT siano lo strumento principale per innovare i business della "old economy".

## 1. SKILL SHORTAGE: C'È ANCORA ?

Negli ultimi due anni un argomento ha "tenuto banco" nel settore delle Tecnologie dell'Informazione e della Comunicazione (ICT): la carenza di specialisti e di competenze di ICT sul mercato, che ha costretto le aziende della domanda e dell'offerta ad inventarsi soluzioni "ponte" per affrontare la crescita tumultuosa dei bisogni e dei fatturati.
Le prime valutazioni nel maggio 2000 indicavano uno skill shortage per l'Italia di oltre 50.000 unità che avrebbero oltrepassato le 100.000 nell'anno 2001.
Questo fenomeno, battezzato "skill shortage", è apparso ufficialmente per la prima volta nel 1999 [5] più o meno in concomitanza col termine "new economy" che evocava una forte discontinuità nello sviluppo delle economie grazie alla tecnologia digitale.
Questo articolo vuole proporre una lettura ampia e critica dei diversi aspetti di questo fenomeno per cercare di dare una risposta alla doppia domanda che oggi emerge dai media e dalla pubblicistica di settore: esiste ancora, in questo quadro di economia mondiale quasi recessiva e soprattutto incerta, un fabbisogno così rilevante di competenze e di specialisti ICT e se sì quali sono le possibili azioni concrete per affrontarlo?
Viene sviluppata una tesi e fornita una metodologia, che non pretende di essere l'unica, ma che ha un interessante punto di forza, quello di essere stata sperimentata ed adottata in un grande Paese come la Gran Bretagna.

## 2. IL "POPOLO" DELLA NET ECONOMY: NASCE NELL'ANNO 2000

Per capire il fenomeno bisogna far parlare un pò i numeri ! Gli occupati nell'ICT sono stati valutati in Italia sino al 1999 a poco più di 441.000 (il 2,1% del totale Italia) in quanto si misurava solo l'occupazione nelle aziende dell'offerta. Nell'anno 2000 la popolazione viene riclassificata, nuovi settori si inseriscono, ed il "popolo" ICT diventa [8] di 1.313.000

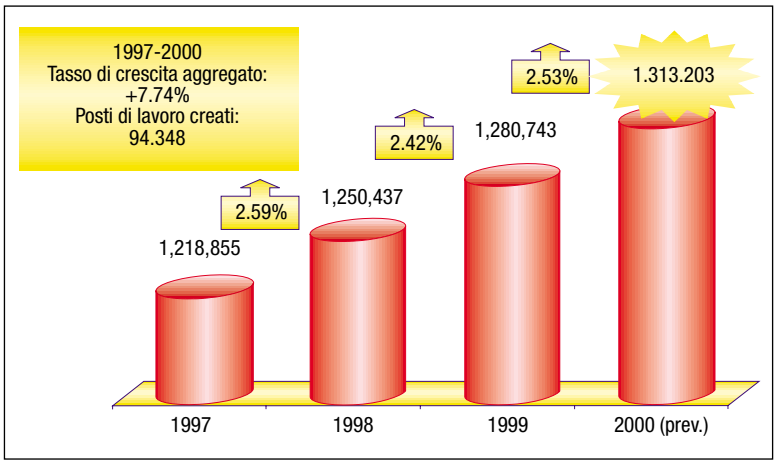



**FIGURA 1**
*L'occupazione nella Net Economy (1997-2000)*
*(Fonte: Federcomin/ NetConsulting su dati ISTAT)*

persone (il 6% degli occupati). Ma non solo: cambia anche nome e diventa *il popolo della Net Economy* (Figura 1).
Questo nuovo modo di vedere il settore fa emergere un dato molto importante: nel periodo 1997-2000 si sono creati oltre 94.000 nuovi posti di lavoro (+7,74%).
Merita capire come sia composta questa popolazione (Figura 2); le novità sono rappresentate dagli addetti dei fornitori di contenuti e servizi di Comunicazione (editoria, radio, televisione, pubblicità), dagli addetti "internet" nelle aziende della domanda, dagli addetti presso le "net company".
Quale futuro potrà avere ? È interessante leggere il rapporto *Employment Outlook 1998-2008 dello US Department of Labour* in cui si vede chiaramente come la Net economy non possa andare oltre il 10% del Pil nei paesi occidentali: in dieci anni l'Europa passerebbe dal 6% all'8% (e l'Italia dal 5% al 7%) e gli USA dall'8% al 10%.
I dati più aggiornati, forniti dall'Osservatorio

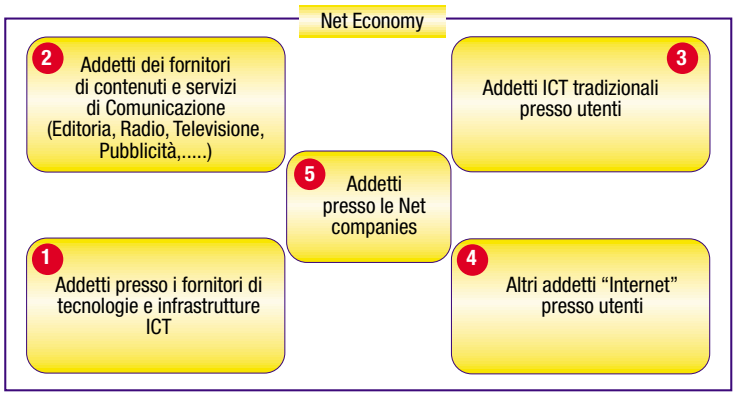



**FIGURA 2**
*L'occupazione nella Net Economy*
*Fonte: Federcomin/ NetConsulting su dati ISTAT*

Federcomin, nell'anno 2001 mostrano alcune novità (Figura 3 e 4).
La novità principale è l'occupazione "collegata" ad Internet ed alle nuove tecnologie che comincia a prendere forma e quantità: nell'anno 2001 viene misurata in almeno 178.000 unità sul totale di 1,4 milioni.

## 3. IL "GAP" DI COMPETENZE È EGUALE PER TUTTI?

La risposta è NO. Sulla base dei numeri, di cui sopra, si possono ipotizzare almeno due scenari per il "popolo" della net economy.
**1.** *Per un numero enorme di addetti, oltre 1,2 milioni*, il principale problema del futuro sarà la gestione del proprio sviluppo professionale per integrare ed acquisire nuove competenze tecniche nelle aree nuove, come Internet, le reti a larga banda, le reti wireless, le nuove modalità di servizio ASP, le opportunità applicative delle nuove tecnologie per lo sviluppo del business (e per il servizio ad imprese e cittadini per le diverse P.A.). Ciò sarà vero sia per la domanda sia per l'offerta pur con esigenze e percorsi diversi.
**2.** *Per gli specialisti* che già oggi operano più direttamente con le nuove tecnologie (almeno 200.000) le sfide sono almeno due: le evoluzioni rapidissime delle nuove tecnologie che comportano monitoraggio ed apprendimento continui, ed i nuovi modelli di business che bisogna "inventare" per sfruttare il potere abilitante delle nuove tecnologie.
Su questi due temi tornerò approfonditamente nel seguito per comporre e capire il quadro di insieme dello *skill shortage* e per identificare strategie adeguate a trattare i diversi fenomeni (vedasi paragrafo 6: "*Un caso emblematico: quale e quanta logistica per l'e-Business*").

## 4. GLI "ALTRI" HANNO PROBLEMI DI "SKILL GAP"?

Ovvero: chi deve affrontare il problema dello *skill gap* e come ? Quando uso il termine "altri" mi viene spontaneo pensare a "tutti gli altri" ....occupati e non occupati ! Può essere molto utile classificare gli "altri" in due grandi categorie:
***a.*** *la categoria degli occupati non "net eco-*

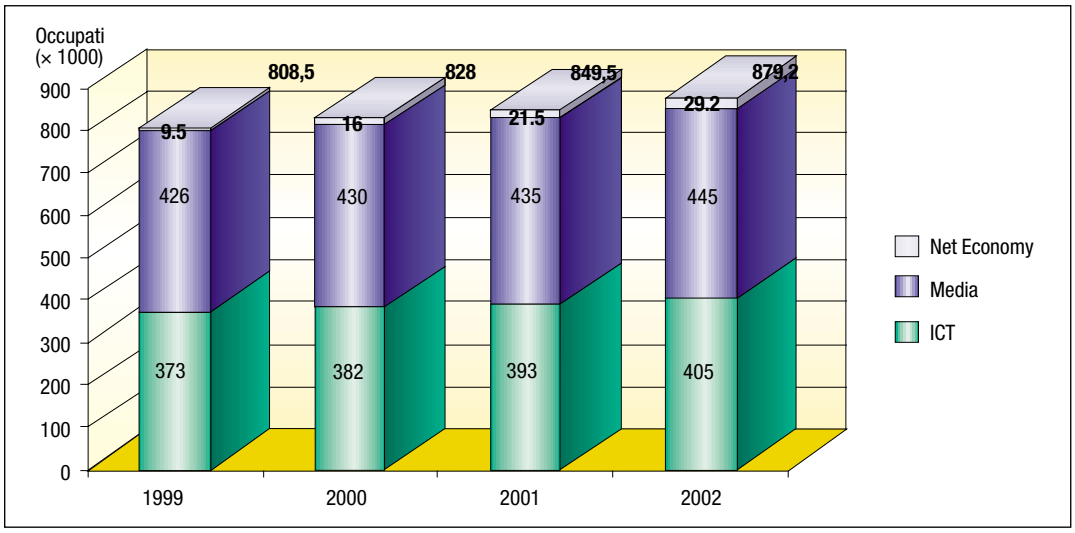



**FIGURA 3**

*La crescita dell'occupazione nelle aziende dell'offerta*

*(Fonte: Federcomin, 2001)*

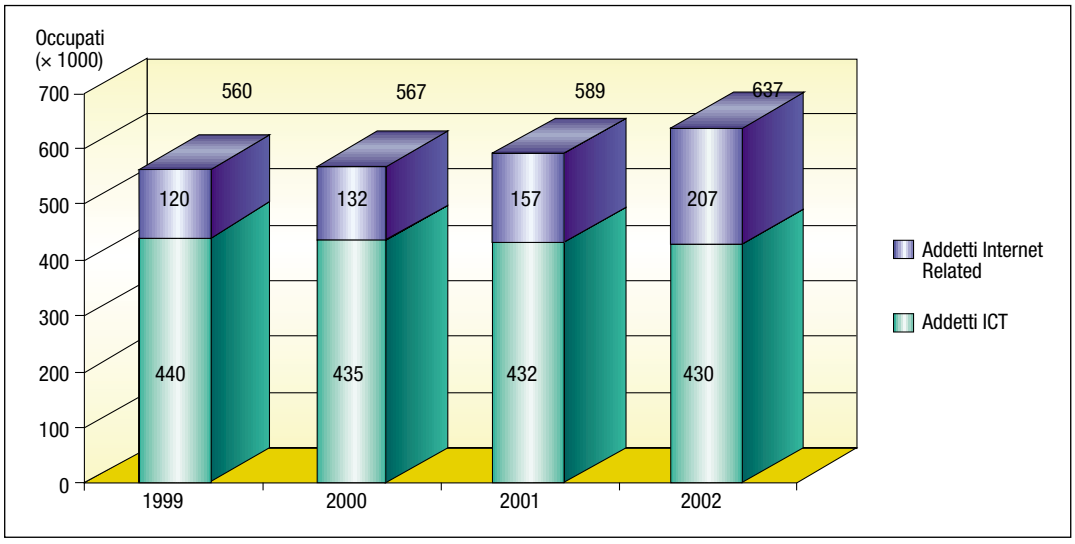



**FIGURA 4**

*La crescita dell'occupazione nelle aziende della domanda*

*(Fonte: Federcomin, 2001)*

*nomy*", che sono... oltre 19 milioni. Qualche volta si è usato il termine "old economy", ma in realtà bisognerebbe dire più semplicemente "economia" o economia tradizionale, non ancora "contaminata" dalle nuove tecnologie della net economy, come Internet e l'e-Business in generale.

***b.*** *la categoria dei consumatori*, cioè tutti noi. Oggi siamo sempre più convinti che lo sviluppo futuro dell'economia sarà spinto dalle nuove tecnologie solo se sapremo dare una risposta convincente su problemi fondamentali: la fiducia tra produttori e consumatori durante le transazioni in rete, la facilità d'uso delle nuove tecnologie, l'accesso generalizzato ed a basso costo per tutti i consumatori.

Lo "*skill gap*", termine che meglio rappresenta l'attuale situazione (più carenze di competenze che di numero di specialisti), riguarda entrambe le categorie ed è quindi un problema ed un fenomeno di massa per i prossimi anni. Tutti devono conoscere e saper usare le nuove tecnologie. *L'alfabetizzazione di base è una necessità vitale per l'intera popolazione*. Le competenze richieste sono di natura ed entità molto differenziate; molti programmi sono già in atto sotto l'egida di aziende, di enti della pubblica amministrazione, del governo centrale e regionale, di associazioni di categoria. C'è un fiorire di idee, dibattiti, iniziative. Anche la Unione Europea e l'OCDE (organizzazione di cooperazione e sviluppo economico) spingono decisamente in questa direzione. È stato coniato il termine "*digital divide*" per evocare pericolose arretratezze culturali che potrebbero mettere in un ghetto le persone

"incompetenti" e addirittura emarginare Paesi che non sappiano imparare la nuova cultura. C'è quindi una prima risposta al quesito iniziale dell'articolo:

> Lo skill gap ICT è una urgenza vera per l'economia italiana nella sua totalità e non solo per la net economy e va affrontata con diversi strumenti a tutto campo, distinguendo tra interventi per le imprese e per la P.A., e quindi per gli occupati, ed interventi sull'intera popolazione di consumatori.

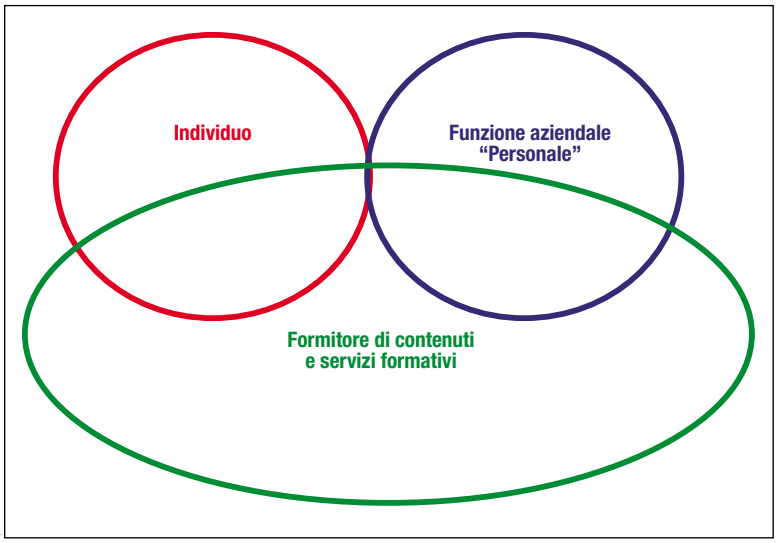
## 5. CON QUALI STRUMENTI SI MISURA E SI RIDUCE LA "CARENZA" DI COMPETENZE?

Ci dobbiamo porre una domanda: stiamo facendo una "ricerca" estesa a tutto il Paese o stiamo cercando una soluzione per la nostra organizzazione? I metodi devono essere diversi nei due casi.

### 5.1. Analisi Paese

Nelle ricerche estese ad un intero Paese il metodo con cui si cerca e si misura lo *skill shortage* è sintetico; si usano cioè campioni di aziende e di persone, e si usano questionari, che possono però rappresentare qualche volta il desiderio piuttosto che la necessità. Con la legge dei grandi numeri e con l'integrazione di altri metodi (Focus Group, Delfi, ad esempio) si converge verso un risultato affidabile. È il miglior metodo che si conosca oggi e quindi possiamo accettarne i risultati per fare diagnosi, e per capire i grandi fenomeni. Non tratteremo nell'articolo questo problema.

**FIGURA 5**
*Il modello. Sviluppo professionale e di carriera*
*(Fonte: Skillmatrix 2001)*




### 5.2. Analisi di una organizzazione

Quando però si deve passare all'azione si devono adottare altri metodi finalizzati ai singoli casi reali, per *confermare la diagnosi e fare la giusta terapia*.
Abbiamo diverse metodologie e casi esemplari di singole Aziende e di "filiere" produttive e di servizio. Da queste diverse situazioni di successo ho estratto tre "strumenti" che, a mio avviso, permettono di dominare il problema e trovare le soluzioni. In primis un modello di riferimento, poi gli standard professionali (utilizzati da oltre 15 anni in UK e da qualche anno in Italia), ed infine l'analisi funzionale.

#### 5.2.1. Un modello di riferimento per gli interventi nelle organizzazioni

Il modello di apprendimento più moderno si presenta come l'intersezione di tre domini (Figura 5):
**i** il dominio *individuale* delle esigenze legate al proprio ruolo, alle proprie competenze, al proprio sviluppo individuale;
**ii** il dominio *aziendale* in particolare quello funzionale del Personale, legato all'obiettivo di sviluppo e di miglioramento delle competenze organizzative;
**iii** il dominio dell'*offerta formativa*, in termini di contenuti e di modalità di apprendimento.
È stato sviluppato un modello dettagliato, per cui all'interno di ogni "circuito", si sono identificate una serie di attività. Un esempio completo per il circuito "individuo" si può vedere nella figura 6 (cortesia di Skillmatrix, 2001); esso mette in evidenza, ad esempio:
- il "profilatore di ruolo";
- il "valutatore a 360 gradi";
- il "Pianificatore di professionalità".

C'è ora una seconda risposta al quesito sul "come" si affronta lo skill shortage:

> Le organizzazioni dovrebbero adottare, per lo sviluppo delle proprie risorse umane, un modello di riferimento basato sull'apprendimento e sulla creazione di conoscenze che integri esigenze ed obiettivi individuali con esigenze ed obiettivi dell'intera organizzazione. Questo modello facilita gli interventi per la riduzione dello "*skill gap*".

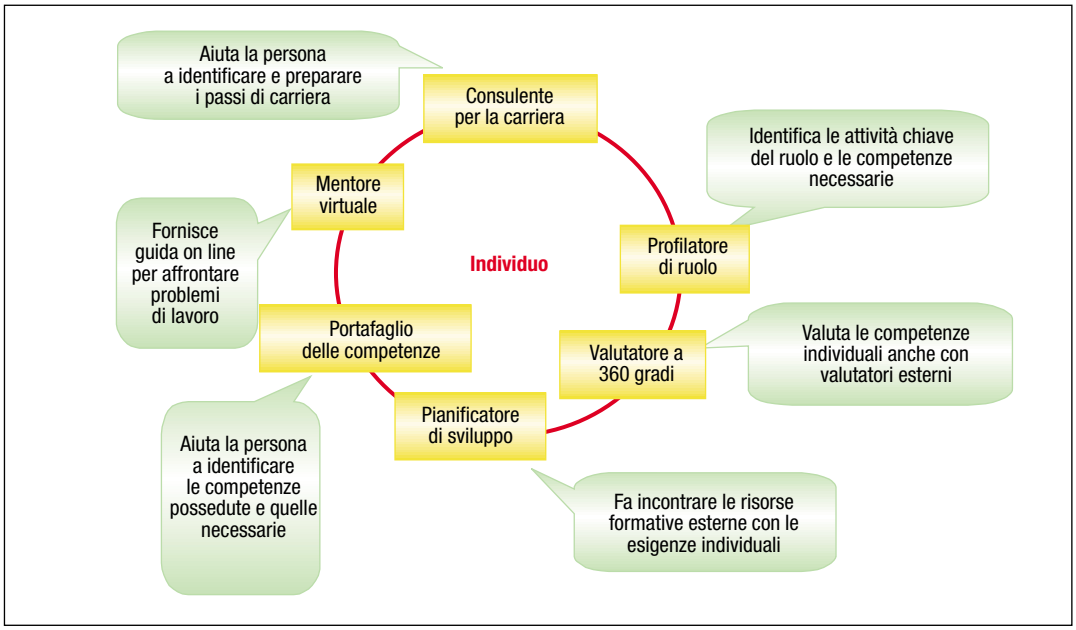



**FIGURA 6**
*Ciclo "individuo"*
*(Fonte: Skillmatrix 2001)*

### 5.2.2. Gli standard professionali

La metodologia più interessante per l'identificazione e lo sviluppo delle competenze, ed anche più sperimentata, è quella degli "standard professionali" [11].
È molto interessante ripercorrere il caso UK. Negli anni '80 in Gran Bretagna il governo, i datori di lavoro ed i sindacati si accordano per creare gli Standard Professionali Nazionali (*National Occupational Standards*): essi definiscono lo standard di prestazione richiesto ai lavoratori a tutti i livelli e per tutti i settori dell'economia.
Detto in altro modo questi standard definiscono le attese di ruolo, e quindi le competenze e conoscenze necessarie per lo svolgimento di ogni attività manageriale e professionale.
In uno studio su di una popolazione di più di 3.000 manager esemplari, la ricerca ha definito l'"obiettivo principale" di tutti i manager: *Raggiungere gli obiettivi della organizzazione e migliorare continuamente i suoi risultati*.
Usando il metodo di *Functional Analysis*, i manager hanno risposto alla domanda ripetuta a vari livelli: *Che cosa bisogna fare per ottenere l'obiettivo principale*? L'analisi ha riconosciuto e analizzato sette funzioni manageriali chiave: gestire attività, gestire risorse, gestire informazioni, gestire persone, gestire ambiente ed energia, gestire la qualità, gestire progetti. Per ognuna delle sette **Funzioni** fondamentali sono stati identificati complessivamente 66 **Processi chiave**. Ad ogni Processo Chiave sono collegati specifici **Sottoprocessi**, **Standard di performance per ogni sottoprocesso, Competenze "soft", Conoscenze ed abilità**.

**UN ESEMPIO**

**Funzione**: *Gestire attività*

**Processo chiave**: Attività per soddisfare il Cliente

**Sottoprocessi**: Concordare i requisiti con i clienti, Pianificare attività per soddisfare i bisogni del cliente, Assicurare la rispondenza di prodotti/servizi ai requisiti

**Standard di performance**: Assicurare supporto alle persone chiave per l'erogazione, Comunicare chiaramente e prontamente con i clienti, Intervenire prontamente di fronte a non conformità

**Competenze "soft"**: Comunicazione, Gestione di team, Influenza sugli altri, Focalizzazione sui risultati

**Conoscenze richieste**: Miglioramento continuo, relazioni con i Clienti, Contesto organizzativo

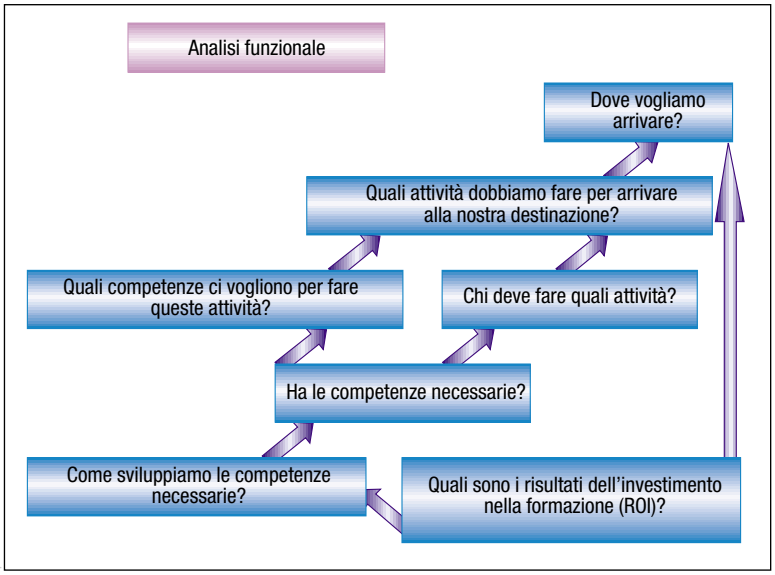
**Abilità richieste**: Affrontare i problemi e sfruttare le opportunità, Controllare la qualità del lavoro e confrontarla con i piani

Attraverso l'esame di quello che fanno i manager di successo, è stato possibile definire *criteri di prestazione* (o fattori critici di successo) per ogni attività. Questi *criteri di prestazione* permettono una valutazione ogget-

tiva dei risultati, e offrono uno strumento valido per la gestione della prestazione stessa (*performance*) e per l'allenamento/addestramento sul campo (*coaching*).
Per ogni attività sono anche state specificate le conoscenze, le capacità e le competenze che un lavoratore deve possedere, o sviluppare, se vuole svolgere il suo compito in modo competente. Queste specifiche permettono un' accurata valutazione delle competenze individuali e la preparazione di programmi di formazione e addestramento *on-the-job* mirati, efficaci ed economici.
In Inghilterra, i lavoratori vengono valutati da valutatori indipendenti, e, se dimostrano di essere competenti, ricevono un Certificato Nazionale Professionale (National Vocational Qualification), spendibile sul mercato del lavoro. Questi standard, quindi, facilitano la mobilità sia interna che esterna.
Una singola azienda o ente può fare la sua analisi funzionale e creare i suoi standard. Però gli standard nazionali sono tutti sviluppati da enti settoriali (per esempio: *Financial Services National Training Organisation, Central Government National Training Organisation*), vengono accettati da tutti le parti interessate, e sono accreditati dal *Department of Education and Skills*.
L'uso degli standard comporta cambiamenti radicali della cultura aziendale. Le agenzie del governo (per esempio l'*Inland Revenue*, l'ufficio delle entrate con 50.000 addetti) usano gli standard per creare una cultura di servizio al cittadino con un livello di servizio analogo a quello fornito dalle migliore aziende private. Le Casse di risparmio e le Assicurazioni mutue hanno trovato negli standard uno strumento che facilita la creazione di una cultura e linguaggio comune quando due o più casse si fondono. Le aziende della Grande Distribuzione Organizzata usano gli standard per ottenere una qualità di prestazione omogenea attraverso manager e commesse in migliaia di negozi geograficamente diffusi.
*Si puo' imparare qualcosa dall'esperienza inglese ed adattare l'approccio delle competenze, basato sul rigore dell'analisi funzionale, al contesto ed al carattere italiano*? La risposta è SI.
Credo che ci sia in Italia, come peraltro in tutta Europa, la forte esigenza di migliorare il profilo di competitività delle imprese e della P.A. centrale e locale, per agire efficacemente in un contesto di globalizzazione. Il modello delle competenze e degli standard professionali collega organicamente le esigenze di sviluppo delle conoscenze individuali e delle organizzazioni con gli strumenti per l'apprendimento.
C'è ora una terza risposta al quesito sul "come" si affronta lo *skill gap*:

Esiste una metodologia, efficace e verificata, con cui analizzare le attività di chi lavora, identificare i "gap" di competenza, assicurando la coerenza con obiettivi e processi dell'Organizzazione. Si chiama "standard professionali" ed è stata sviluppata ed applicata con successo in UK negli anni '90.

**FIGURA 7**
*(Skillmatrix, 2001)*




### 5.2.3. L'ANALISI FUNZIONALE

Una metodologia molto efficace per capire e collegare tra di loro i diversi "componenti" di una organizzazione è l'"analisi funzionale". Essa permette di collegare in un quadro comprensibile ed integrato sei componenti:
1. la missione aziendale;
2. le attività chiave che permettono di realizzare la missione;
3. i ruoli organizzativi;
4. le competenze necessarie sottese ai ruoli;
5. il "gap" di competenze per ogni persona a cui è affidato il ruolo;
6. i percorsi di sviluppo per ciascun individuo per superare i "gap".

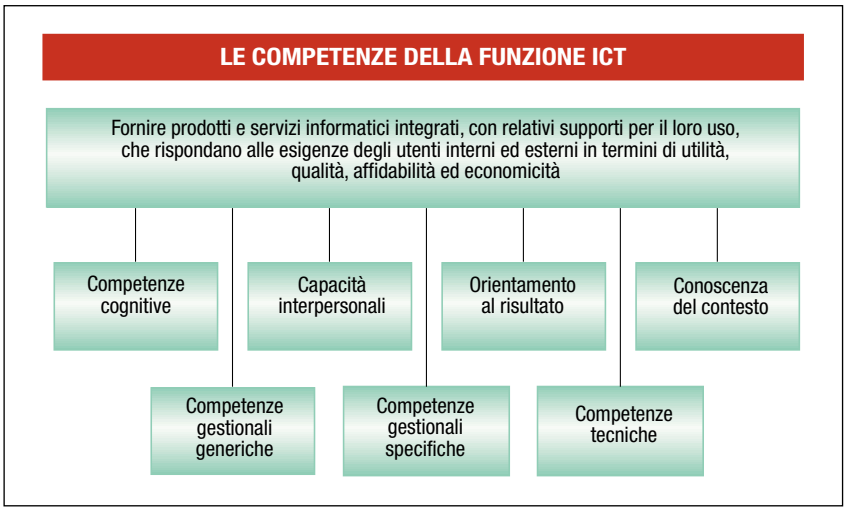
Nella figura 7 è esemplificata l'applicazione di questo metodo.

***Un esempio di applicazione del modello degli standard professionali e della "analisi funzionale" ad una intera funzione aziendale ICT.***

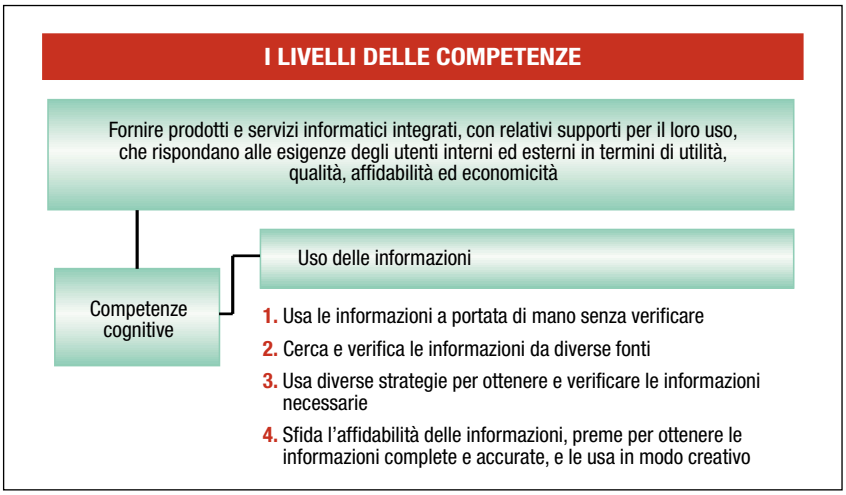
*Dove vogliamo arrivare?*


*Quali attività dobbiamo fare per arrivare alla nostra destinazione?*

Quali competenze ci vogliono per fare queste attività?




*A quale livello deve essere la specifica competenza?*
*1 basilare, 4 massimo*




## 5.3. La patente informatica ECDL

La Unione Europea, attraverso il CEPIS (*Council of European Professional Informatics Societies*), l'ente che riunisce le Associazioni europee di informatica, ha promosso uno strumento per l'alfabetizzazione di base degli utilizzatori delle tecnologie dell'informazione. Si chiama "***Patente europea di guida del computer***" (ECDL: *European Computer Driving Licence*). In Italia l'AICA [2] è l'ente nazionale di certificazione del programma ECDL.
È stato cioè definito uno standard per misurare le competenze di utilizzo dell'ICT che ha creato anche i percorsi formativi per raggiungere il livello base di conoscenze e di capacità. Ma che cosa significa realmente saper usare il computer? È sufficiente scorrere la lista degli esami da sostenere per ottenere la "patente" (vedasi scheda allegata) per rendersi conto che l'ECDL è una piattaforma di base di competenze, assolutamente necessaria per poter "navigare" nelle applicazioni informatiche e nella rete Internet, e poter usare efficacemente i più comuni strumenti di informatica individuale.
In base a un *protocollo di intesa con l'AICA*, il Ministero della Pubblica Istruzione ha adotta-

### COME SI OTTIENE LA PATENTE EUROPEA DEL COMPUTER?

Il candidato deve acquistare da un qualsiasi Centro accreditato (*Test Center o Test Point*) una tessera (*Skills Card*) su cui verranno via via registrati gli esami superati.
Gli esami sono in totale sette, di cui uno teorico mentre gli altri sono costituiti da test pratici. Il livello dei test è volutamente semplice, ma sufficiente per accertare se il candidato sa usare il computer nelle applicazioni standard di uso quotidiano. Più precisamente, sono previsti i seguenti moduli:
1 - Concetti teorici di base (*Basic concepts*)
2 - Uso del computer e gestione dei file (*Files management*)
3 - Elaborazione testi (*Word processing*)
4 - Foglio elettronico (*Spreadsheet*)
5 - Basi di dati (*Databases*)
6 - Strumenti di presentazione (*Presentation*)
7 - Reti informatiche (*Information networks*)
Ogni esame può essere sostenuto presso un qualsiasi Centro accreditato in Italia o all'estero. Il candidato non è, cioè, obbligato a sostenere tutti gli esami presso la stessa sede e inoltre può scaglionarli nel tempo (la tessera ha una validità di tre anni).
Quando ha superato tutti gli esami, egli riceve la patente (diploma) da parte dell'ente nazionale autorizzato ad emetterla (in Italia, l'AICA). Ogni esame può essere sostenuto presso un qualsiasi Centro accreditato in Italia o all'estero. Il candidato non è, cioè, obbligato a sostenere tutti gli esami presso la stessa sede e inoltre può scaglionarli nel tempo (la tessera ha una validità di tre anni).
A maggio 2001 AICA ha introdotto un nuovo sistema, ALICE, che automatizza in modo integrale gli esami che abilitano al rilascio del diploma ECDL. Per maggiori informazioni su ALICE si veda la pagina
http://www.aicanet.it/alice/alice.htm
informazioni generali: http://www.aicanet.it/ecdl.htm
materiale didattico: http://www.aicanet.it/materiale_didattico.htm

to ECDL come standard per la certificazione delle competenze informatiche nella scuola. Di conseguenza la patente europea del computer è accettata, senza problemi, come credito formativo negli esami di stato per il diploma di maturità.

## 6. UN CASO EMBLEMATICO: QUALE E QUANTA LOGISTICA PER L'E-BUSINESS

Un gruppo di lavoro dell'AILOG ([3] Associazione italiana di logistica e di supply chain management) ha analizzato nel 2001 l'interazione tra le tecnologie ICT, in particolare l'e-Business, e la Logistica giungendo a conclusioni che sono molto interessanti per il ragionamento che stiamo conducendo sullo "*skill shortage*", che diventa sempre più "*skill gap*". Sono stati estratti nove messaggi logistici, ossia nove sfide che le nuove tecnologie da un lato "creano" perché generano cambiamento e dall'altro "risolvono" offrendo soluzioni.
Il messaggio centrale è molto semplice e diretto: **e-Business è soprattutto Logistica e sempre più Logistica**. La Logistica *non è solo lo strumento che abilita il business sulla rete Internet*, è qualcosa di più, ossia sempre di più la Logistica è il modello di business, è la business idea. Si evidenzia la centralità della competenza Logistica ai diversi livelli di progettazione e gestione: nuovi modelli di impresa virtuali, nuovi modelli di business logistico-centrici, creazione del valore nella supply chain, gestione a distanza di processi di codesign e comakership.
Logistica e ICT insieme stanno innescando dei cicli di sviluppo importanti.
**Si devono creare nuove competenze per progettare e governare i sistemi di e-Business e in generale l'innovazione**.
Come possiamo leggere questi messaggi? Secondo me la lettura principale è che i confini delle competenze, nello specifico tra Logistica ed ICT, è sempre meno definibile; sempre di più emerge ***la competenza della "innovazione" che è trasversale***. Queste capacità di "trasversalità", cioè di lavorare insieme, di integrazione di conoscenze, diventano essenziali per trovare idee e soluzioni ai problemi; lo "skill gap" sta diventando quasi un fatto costitutivo della sfida della integrazione

delle nuove tecnologie nell'economia e nei modelli correnti di business.
La quarta risposta alla domanda iniziale dell'articolo potrebbe proprio essere questa:

> Se vogliamo veramente sfruttare il potere abilitante delle nuove tecnologie bisogna sviluppare nuove competenze, non solo tecniche ma anche manageriali, nuovi profili professionali, nuove modalità di apprendimento.

## 7. CONCLUSIONI

Mi piace citare Gian Filippo Cuneo [10] che scrive l'11 aprile 2001 nella sua mailing "Pensieri laterali":
*"Internet è qui per restare; chi prende sollievo da qualche fallimento di impresa virtuale vuole solo allontanare da se' la fastidiosa inevitabilità del cambiamento che tocca tutte le imprese. Accettare l'era di internet, evitare i pericoli di una risposta ritardata e svogliata, e coglierne le opportunità è prima di tutto una questione di schemi mentali di riferimento; le imprese italiane hanno normalmente schemi vecchi, inadatti all'era di Internet.....L'imprenditore che non ha ancora capito il fenomeno Internet pensa che l'on-line sia un fatto eccezionale, che l'impresa debba essere integrata, che sia importante produrre prodotti invece che servizi, che il servizio debba essere fatto da umani e non da computer, che sia sufficiente una segmentazione della clientela approssimativa, che controllare il capitale sia sufficiente a gestire imprese basate sulla conoscenza, che esista un mercato domestico, e che basti fare profitti per vincere la concorrenza e mantenere il valore dell'impresa. L'errore primigenio sta nel pensare che il mondo di ieri sia anche quello di domani; purtroppo, però, il futuro non e' più quello che era."*
***Lo skill gap piuttosto che lo skill shortage è la vera sfida nei prossimi anni.***
Lo *skill gap*, parafrasando scherzosamente la citazione di cui sopra, "*è qui per restare ed è prima di tutto una questione di schemi mentali di riferimento*".
Gli standard professionali, le metodologie per rilevare le attività e le competenze, e le stesse tecnologie ICT, ci possono peraltro aiutare moltissimo ad impostare un buon piano d'azione per affrontare questa "emergenza, che sta diventando continua".
La principale raccomandazione è quella di dotarsi di un buon schema di riferimento, cioè di un modello che renda chiari ed espliciti i collegamenti tra i bisogni ed i desideri degli individui e gli obiettivi e gli schemi di lavoro delle organizzazioni in cui questi individui operano. Il segreto potrebbe proprio essere quello di "vedere" ed "agire" in modo integrato: integrazione tra individui ed organizzazione, integrazione tra discipline.

## Bibliografia


[1] AAVV: *Harvard Business Review on Knowledge Management*. Harvard Business School Press, 1998.

[2] Aica: *La patente informatica ECDL*, http://www.aicanet.it/ecdlwhat.htm

[3] Ailog: *Gruppo di lavoro Logistica&e-Business, anno 2001*, www.ailog.it

[4] Anasin-Federcomin: *ict-job, Internet*, http://www.ict-job.it/

[5] Assinform: *Rapporto sull'informatica e le Telecomunicazioni, anni 1998, 1999, 2000, 2001*, www.assinform.it

[6] Assinform: *Rapporto 2001 sull'occupazione nel settore dell'Informatica e delle telecomunicazioni in Italia*, www.formazioneict.assinform.it

[7] Charles Leadbeater: *Living on thin air*. Penguin Books, 2000.

[8] Federcomin, Rapporto: *Occupazione e formazione nella Net economy*, 2000, 2001, www.federcomin.it

[9] FidaInform: rivista n. 22, settembre 2001, intervista a Giovanni Degli Antoni, Skill shortage, un falso problema, www.fidainform.org

[10] Gianfilippo Cuneo: *Pensieri laterali*,11 Aprile 2001, pensieri_laterali@it.buongiorno.com

[11] Trevor Boutall: *The good Manager's guide, MCI*, London, 1997.



RENZO PROVEDEL, ingegnere, ha sviluppato competenze di management in grandi aziende (Fiat, Iveco, ILVA) nei settori Informatica, Telecomunicazioni, Organizzazione e Logistica, ricoprendo responsabilità a livello Corporate, divisionale ed internazionale. Oggi è Imprenditore della net economy, titolare di un'azienda di servizi alle imprese per l'applicazione delle tecnologie ICT in processi di innovazione e di riposizionamento strategico. Presidente di FareImpresa, consigliere di AICA, di FidaInform, di AILOG.
E-mail: provedel@fareimpresa.net